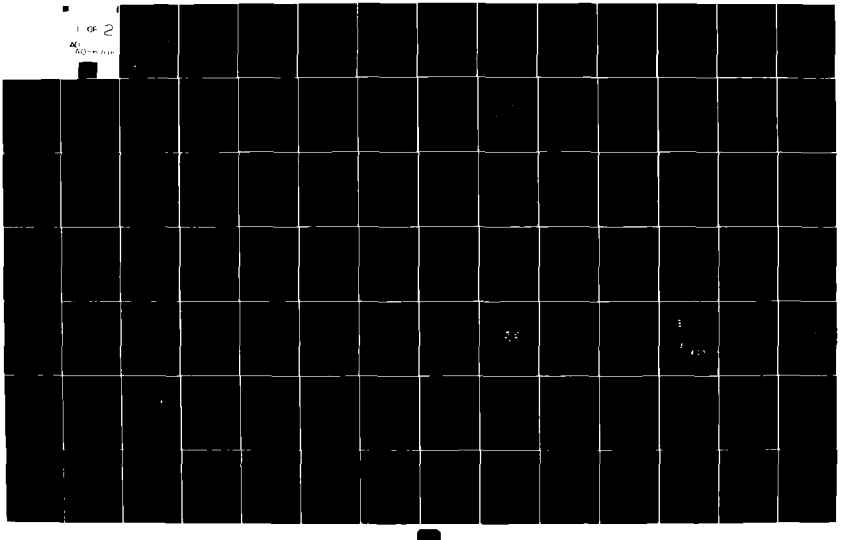
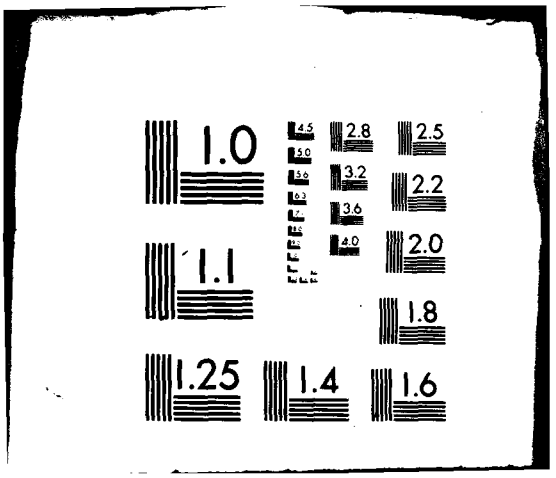
AD-A086 708 GENERAL ACCOUNTING OFFICE WASHINGTON DC PROGRAM ANAL--ETC F/G 15/3
MODELS, DATA, AND WAR: A CRITIQUE OF THE FOUNDATION FOR DEFENSE--ETC(U)
MAR 80
UNCLASSIFIED GAO/PAD-80-21 NL




1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6

BY THE COMPTROLLER GENERAL

Report To The Congress

OF THE UNITED STATES

Models, Data, And War: A Critique Of The Foundation For Defense Analyses.

Under such names as operations research, computer modeling, and cost-effectiveness analysis, quantitative tools and methods have come to play a prominent role in the analysis of public policy issues.

Weapon systems costing hundreds of millions or billions of dollars, composition of future forces, and other defense planning and decisionmaking often are justified in part, or supported, by quantitative studies. DOD estimates that the annual cost of such studies is about a quarter of a billion dollars.

This report's recommendations are intended to make Defense studies and analyses more responsive to the needs of high-echelon decisionmakers.

COMPTROLLER GENERAL OF THE UNITED STATES
WASHINGTON, D.C. 20548

B-197094

To the President of the Senate and the
Speaker of the House of Representatives

This report critiques the management and use of quantitative methodology in the analysis of public policy issues, focusing on the inherent limits of the methodology as a tool for Defense Decision, and the essential role of human judgment in any such analysis.

Copies of the report are being sent to the Director, Office of Management and Budget; the Secretary of Defense; the Chairman, Joint Chiefs of Staff; and to the Secretaries of the Army, Navy, and Air Force.

Comptroller General
of the United States

COMPTROLLER GENERAL'S REPORT TO THE CONGRESS

MODELS, DATA, AND WAR: A CRITIQUE OF THE FOUNDATION FOR DEFENSE ANALYSES

D I G E S T

A prominent feature of modern Government is the extent to which the executive branch has institutionalized quantitative methodology (cost-effectiveness analysis, computer modeling, etc.) as an aspect of budgeting and decisionmaking. Proponents term this "scientific management," and view it as a salutary extension of the "objective" tools of science and mathematics. This argument has substantial merit, but it obscures the fact that quantitative methodology has considerable potential in both scientific (or "objective") and "subjective" applications. The difference, whether an application is based on scientific fact or "quantified judgment," has obvious importance in the context of decisionmaking.

This report examines the nature of quantitative methods (cost-effectiveness analysis, computer modeling, etc.) and some of the problems in their use for the analysis of public policy issues.

A major contention of this report is that quantitative techniques have considerable potential as an aid in the analysis of public policy issues, but that this potential is impaired by the current design and management of quantitative tools. Improving these tools--providing better information for tomorrow's decisionmakers--is the theme of this report and the purpose of its recommendations.

GAO's findings and recommendations should be of interest to all Federal agencies who do, use, or rely on quantitative analysis. They should be of special interest to Federal decisionmakers whose demands for better information create the pressures for improvement.

The report concentrates on the Department of Defense's effort to examine conventional ground and tactical air force requirements by mathematical-statistical means. A combat or campaign model is one part of this activity.

Expert judgment, empirical data, and a quantitative theory of combat are the other necessary parts. The discussion entails gross simplifications about the range and scope of models, issues, and organizational entities involved in Defense decisionmaking. Inclusion of real world complexities would not change the basic message.

A ROLE FOR QUANTITATIVE ANALYSIS

From a scientific point of view, the present "understanding of war"--insofar as the effectiveness of conventional military forces is concerned--is in a relatively primitive state. Basic research aimed at understanding the fundamentals of combat is needed, but quantitative or numerical techniques have not been systematically applied to achieve these discoveries.

What is the effectiveness of a weapon system?

How can its effectiveness be measured?

How can this be linked to the broader objective of determining the most effective mix of ground and tactical air forces that can be bought and maintained for, say $30 billion per year?

Defense decisions--whether based on military judgments or sophisticated economic techniques --are critically dependent on the knowledge of what a military force can reasonably be expected to do. That knowledge is essential, not only to permit realistic force comparisons, but also for gauging the individual contributions of new weapons and tactical concepts.

MATTERS FOR CONSIDERATION BY DEFENSE DECISIONMAKERS

The quantitative method offers the opportunity to bring together the "best" of science and considered judgment. That it has aided Defense decisionmaking is without question. But if the findings of this and earlier studies are representative, its full potential has not been achieved. To do so, Defense decisionmakers must act on the premises that:

--Quantitative decisionmaking is beneficial only when it embodies, rather than replaces, expert judgment and "objective" fact.

--Analyses may give the appearance of scientific work but may not have been subjected to the normal evaluative standards of science.

--The theory and supporting data may not equal the quality of the analytic tool.

--The assumptions and limitations of the analysis must be made a part of the study report.

RECOMMENDATIONS TO THE SECRETARY OF DEFENSE

The Secretary of Defense should reassess the adequacy of current practices in the management and use of policy assisting models employed in Defense Decision. This should include identifying needed corrective measures and insuring that such models are used to enhance and extend the decisionmaker's judgment. The Secretary should also develop procedures to enhance the contribution of policy assisting models to open explicit analysis in key areas of policy, strategy, and force planning. (See p. 80.)

Further, the Secretary should require the Chairman, Joint Chiefs of Staff, to review current procedures for safeguarding and strengthening the empirical-theoretical foundation underlying the representation of combat in Defense studies. And, as warranted by that review, the Chairman should be required to prepare plans and recommendations which would enable the Organization of the Joint Chiefs of Staff to serve as the Defense Establishment's principal analytic adviser on matters pertaining to the phenomenology of combat. (See pp. 104-105.)

AGENCY COMMENTS AND EVALUATION

The Department of Defense believes that its current structure and directives provide for satisfactory management of Defense analyses. The Department's detailed comments and GAO's evaluation are included as appendix IV.

GAO's recommendations are directed toward establishing a consistent analytical framework to support high-echelon Defense decisionmaking. These, or similar, actions are necessary to alleviate what GAO believes is a continuing weakness in the design for Defense analyses.

RECOMMENDATIONS TO THE CONGRESS

This report focuses on the inherent limits of quantitative methodology as a tool for Defense Decision, and the essential role of human judgment in any such analysis. But such limitations are not restricted to defense issues. They also affect mission budgeting, risk assessment, the evaluation of social programs, and the like. Congressional concern with the quality of these analyses and their supporting tools is a matter of record.

When reviewing quantitative studies or exercising its oversight authority, the Congress should require

--an open, explicit understanding of the assumptions underlying a study's conclusions, or knowledge of the identity of the decisionmaker(s) involved in the study, their background experience and institutional affiliation(s); and the extent to which the model(s) used in the study have been appraised, possibly in the form of an explicit statement accompanying the results of the study.

--agencies and departments to report on how current decisionmaking tools are being managed, and what is being done to improve the tools and provide better answers to difficult public policy questions in the future.

When considering defense acquisition requests and cost-effectiveness analyses, the Congress should inquire how a particular program or weapon system contributes to the overall force level analyses. (See the suboptimization example on p. 45.) In this way, the Congress can ascertain the links between a weapon's effectiveness and the quantity required for a defense mission.

These are all matters which have considerable bearing on the combat readiness of the Armed Forces as well as on defense expenditures.

Contents

Page

TABLES

FIGURES

ABBREVIATIONS

APC	armored personnel carrier
ATLAS	A Tactical, Logistical, and Air Simulation
CAS	close air support
CCTC	Command and Control Technical Center
CEM	CONAF Evaluation Model
COEA	Cost and Operational Effectiveness Analysis
DARPA	Defense Advanced Research Projects Agency
DOD	Department of Defense
FEBA	forward edge of the battle area
GACAM	Ground-Air Campaign Model
GAO	General Accounting Office
GNP	Gross National Product
IDAGAM	Institute for Defense Analyses Ground-Air Model
JCS	Joint Chiefs of Staff
MOE	measure of effectiveness
NATO	North Atlantic Treaty Organization
NTC	National Training Center
OSD	Office of the Secretary of Defense
PIES	Project Independence Evaluation System
POM	Program Objective Memorandum
PPBS	Planning-Programming-Budgeting System
R&D	research and development
SAGA	Studies, Analysis, and Gaming Agency
SECDEF	Secretary of Defense
TRIM	Transfer Income Model

CHAPTER 1

SCOPE OF REPORT

A prominent feature of modern Government is the extent to which the executive branch has institutionalized quantitative methodology (cost-effectiveness analysis, computer modeling, etc.) as an aspect of budgeting and decisionmaking. Proponents term this "scientific management," and view it as a salutary extension of the "objective" tools of science and mathematics. This argument has substantial merit, but it obscures the fact that quantitative methodology has considerable potential in both scientific (or "objective") and "subjective" applications. The difference, whether an application is based on scientific fact or "quantified judgment," has obvious importance in the context of decisionmaking.

This report critiques the role of quantitative methodology in the analysis of public policy issues. It focuses on the inherent limits of the methodology as a tool for Defense Decision, and the essential role of human judgment in any such analysis. The emphasis is methodological, and deals with the quality of the tools provided for the analysis of complex policy problems. For this reason, some of the criticisms made here may appear to be overstatements when related to simpler problems and less ambitious objectives.

For the foreseeable future, a major dilemma for the United States will be the need to reconcile demands brought on by a numerically impressive military threat with pressures for nonmilitary programs, while allocating defense dollars between expensive systems in smaller quantities and cheaper systems in larger quantities. The task is formidable, but not unmanageable. Simply stated, these and many other policy issues must be decided by the human mind. The "promise" of quantitative methodology is to enhance and extend these judgments.

A major contention of this report is that quantitative techniques have considerable potential as an aid in the analysis of public policy issues, but that this potential is impaired by the current design and management of quantitative tools. Improving these tools--providing better information for tomorrow's decisionmakers--is the theme of this report and the purpose of our recommendations. The critical problems, in our view, are not technological; they are institutional, a point not widely recognized. Our discussion should be of legitimate interest to all Federal agencies who do, use, or rely on this type of analysis. It should be of special interest to Federal decisionmakers, for they are

the demanders who, in large part, create the institutional pressures for improvement.

Our presentation concentrates on the Department of Defense's effort to examine conventional military force requirements by the techniques of quantitative methodology. A combat or campaign model is one part of this activity. Expert judgment, empirical data, and a quantitative theory of combat are the other necessary parts. This report treats the four subjects. Our work is drawn from:

--A comprehensive review of over 900 technical reports and articles (classified and unclassified) which span the period from World War II to date, most published or funded by DOD. The material presented here is unclassified.

--Interviews with senior officials in the Office of the Secretary of Defense, the Organization of the Joint Chiefs of Staff, and the Services; and with selected civilian and military professional staff members in the various DOD studies and analysis activities.

--Interviews with officials from selected Federal Contract Research Centers and private research and consulting organizations, including:

CACI, Incorporated	Institute for Defense Analyses
Center for Naval Analyses	Naval Postgraduate School
General Research Corporation	The RAND Corporation
Historical Evaluation and Research Organization	SRI International
	Vector Research, Incorporated

The ideas and insights obtained during these interviews have materially contributed to our work. The views and opinions expressed here are, of course, our own.

CHAPTER 2

THE FOUNDATION FOR DECISION:

MATHEMATICS--OR SCIENCE AND EXPERT JUDGMENT?

The past two decades have seen a tremendous expansion in the role and influence of computer models as a "policy assisting" device--first in the analysis of national security issues, and now, for the analysis of various problems arising in the affairs of Government. That the use of such models can be fruitful and rewarding is without question; but to view them as an extension of the "objective" tools of science and mathematics is a serious mistake. The "policy assisting" models differ from their engineering-accounting cohorts in ways that are subtle, and yet nontrivial. They will easily lead an unsuspecting decisionmaker down the garden path.

POLICY ASSISTING MODELS DIFFER FROM THEIR ENGINEERING-ACCOUNTING COHORTS

In what follows, the term "policy assisting model" connotes a computer model that (1) is used for the systematic examination or analysis of squishy problems 1/ (i.e., problems without a well-defined mathematical representation); and (2) is intended to influence high levels of Government decisionmaking. It includes models that are used by agency and executive branch officials, as well as models that influence congressional debate and action. Above all, the term describes a model that deals with questions beyond the purview of rigorous scientific deduction.

The problems are beyond science

Many of the problems encountered in the classical sciences, engineering, or accounting, are rigorously quantifiable. Their structure is well-understood, and their mathematical formulas provide a clear-cut representation of the "real world" problem. A squishy problem, on the other hand, may be given a mathematical form that looks like an unambiguous representation of the real world problem; but the appearance is only superficial, and evaporates rapidly when probed to any great extent. Thus, for example, the questions: "What military forces are needed to defend NATO?";

1/We borrow this term from Ralph Strauch. The genesis for much of our presentation may be found in his cogent discussion of the squishy problem and its influence on policy analysis and quantitative decisionmaking. See Strauch [81].

"How will the Social Security Amendments affect capital formation?"; and "How much grain will be available for export in the coming year?" are squishy problems. The mathematical form given to such problems is inherently judgmental, and the conclusions reached are inextricably tied to those judgments. Different analysts, with apparently identical knowledge of a real world problem, may develop plausible formulations that lead to very different conclusions--none of which are verifiable or refutable.

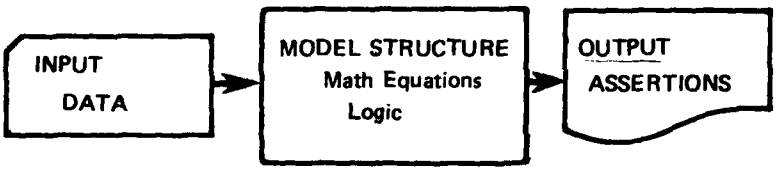
As a rule, problems with any degree of behavioral content, and those that border on the threshold of science, tend to be squishy. The more central these features become to the real world issue (as when NATO issues depend on the tenacity of combat forces, or when agricultural issues depend on long-range weather conditions), the squishier the problem is likely to be. Such problems are often central to the analysis of national defense, food, energy, social welfare, and other issues.

The models only manipulate numbers

To begin understanding what is at stake, one must understand that a computer model is a "purely quantitative" device --simply an analytic abstraction of a more complex real world problem or phenomenon. The model's structure--specified by equations and other forms of quantitative logic--merely transforms input data into numerical assertions that represent real world outcomes (see figure 2.1).

FIGURE 2.1

THE COMPUTER MODEL



Computers can manipulate impressive quantities of variables, and their power is sufficient to permit the building of large and detailed models supposedly depicting complex policy settings. But models perform only mathematical analysis--and mathematics (a logical language) is one thing;

science (an understanding of the real world), quite another. Mathematical validity is a necessary--but not always a sufficient--condition for certifying a model's conclusions. Herbert Simon has emphasized the importance of this distinction in remarking,

> "Mathematical social science is first and foremost social science. If it is bad social science (empirically false), the fact that it is good mathematics (i.e., logically consistent), should provide little comfort." 1/

The solutions are only judgments

> "A model, composed of variables and parameters in specific configurations, is a form of theory. Which variables and parameters are included represents a fundamental theoretical choice, and how those entities are configured is a concrete theoretical statement." 2/

The formulation of a computer model--conceiving a mathematical representation of the real world--can be a highly intuitive process, so much so that the policy assisting models give lie to the notion that knowledge produced by the application of quantitative methodology is "objective," whereas professional judgment is subjective. Concisely,

- Squishy problems are like the issues that are widely debated in the literature, but on which experts disagree.
  - -- Neither science, nor empirical evidence, nor a consensus of expert opinion, provide a mathematical form to describe their structure, or predict their behavior.
- Policy assisting models, therefore, do not produce "objective," or verifiable, solutions.
  - -- The model's structure may be an admixture of science, empirical research, and theory; but it is the builder's judgment that supplements and melds these things into a concrete theoretical statement.

1/See Herbert A. Simon, in Lazarsfeld [60], p. 388.

2/See Brewer and Hall [13], p. 12.

The message is simple and clear. Policy assisting models --inescapably grounded in the builder's judgment, and producing unverifiable results--are used in the analysis of some of the most important problems confronting Government decisionmakers.

To expect these models to produce "objective," scientifically valid results is no more reasonable than to expect that a particular brush will produce fine paintings, or a particular knife fine carvings. Policy assisting models are intended to be used as an extension of, rather than a replacement for, human judgment. They are "givers of insight" (even under the conditions defined by the model) rather than producers of scientifically valid results. In fact, they are used to analyze problems for which there is no scientific, or "objective" answer.

Policy assisting models are one of the most significant decisionmaking tools of our day. But the services they provide, and the management attention they need, are very different than what we have come to expect with their engineering-accounting (or rigorously quantifiable) cohorts. They should be looked upon not as a "provider of solutions," but rather as a framework which permits science and the judgment of experts in numerous subfields to be brought together--made explicit--and utilized to enhance and extend a decisionmaker's judgment. This is their aim and opportunity. To accept any other view uncritically is both naive and dangerous.

ANALYSIS OF A RIGOROUSLY QUANTIFIABLE PROBLEM

The simplest applications of quantitative methodology, from the viewpoint of the problem/model relationships involved, are those in which the structure and logic of the real world problem are the same as the structure and logic of the model being used. These "rigorously quantifiable" problems arise, for example, in the analysis of well-understood systems that obey accepted scientific laws, accounting procedures, and the like.

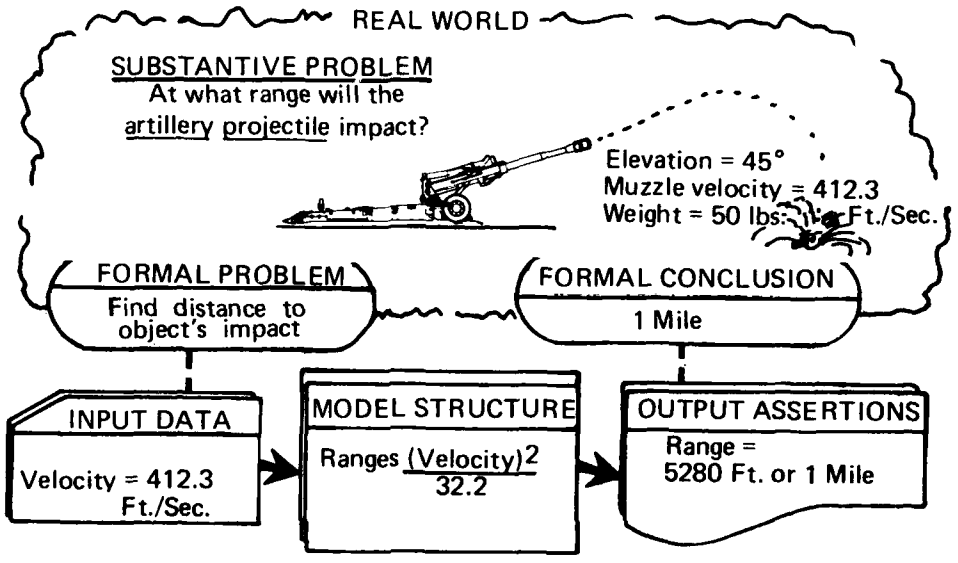
A scientific formulation

The calculation of an artillery projectile's range is an example of a rigorously quantifiable problem (see figure 2.2). The analysis has two levels, a substantive (real world) level and a mathematical level. The analyst begins with the substantive problem, develops a mathematical model of the problem, analyzes the model to produce mathematical results,

and then projects the results back up to the real world level as substantive conclusions. 1/

Figure 2.2

Analysis of a Rigorously Quantifiable Problem



The analysis is methodologically pleasing because Newtonian physics (as a mathematical surrogate or model) captures the essence of the substantive problem. The coincidence of structure between model and problem insures that the model's output assertions will also be statements about the problem.

To find the artillery projectile's range, the analyst enters particularizing data--say, the gun's elevation and muzzle velocity--and then accepts the model's output assertion as a substantive conclusion about the real range of the real projectile. This can be done without ever consciously thinking about the fact that Newtonian physics is

---

1/The Newtonian formula in figure 2.2 is merely illustrative. More complex versions--incorporating the effects of wind, atmospheric density, projectile shape, etc.--are commonly used for artillery calculations.

not the real world process at all, but simply a mathematical abstraction that serves as a convenient surrogate. 1/

An "objective" solution

Notice, however, that the premises defining the model are not simply mathematical premises--they are also assumptions about the behavior of the real world. Thus, confidence in the model's output assertion depends on both: (1) the internal logical validity of the mathematical analysis, and (2) the empirical validity of the linkages between the model and the substantive problem. In problems dealing with systems obeying well-understood and accepted scientific laws (or accounting procedures, etc.), the empirical validation is provided by the fact that the models used represent those accepted laws (or procedures). In other cases, the validity of treating the model as a surrogate for the real world problem must be demonstrated--through statistical experimentation, long empirical experience with the process, etc. If these mathematical and empirical validity criteria are met, then the conclusions reached are "objective" in the sense of being grounded in rigorous logic and objective fact.

In point of fact, the rigorously quantifiable models are "free-standing"; their validity can be decided on the basis of structure alone, without reference to the competence or judgment of their creators. When properly documented, such models can be verified or refuted by an independent critic--solely on the grounds of the mathematical analysis within the model and the empirical connection between the model and the substantive problem.

To recapitulate, rigorously quantifiable models are

- verifiable by an independent critic; and
- "objective" in solving real world problems (implying that a validated model can be used to replace human judgment).

1/Here, we emphasize the criterion of accurate, rather than exact, predictions. (Few things can be specified absolutely. Even the decimal equivalents of fractions such as one-third and one-sixth are approximations. Their "accuracy" can be assessed only in relation to a defined level of "precision.") In this report, we use "rigorously quantifiable" to mean that the model's assertions are accurate for the context and precision of a substantive conclusion.

ANALYSIS OF A SQUISHY PROBLEM

If all problems were rigorously quantifiable, there would be no need for concern about the limits of quantitative methodology, or about applications not clearly justified by theory. But that is not the case. Government decisionmakers deal with a spectrum of problems--ranging from rigorously quantifiable at one end, to highly squishy problems of dubious quantifiability at the other end. Significant policy problems, in particular, tend to lie much nearer the squishy end.

A judgmental formulation

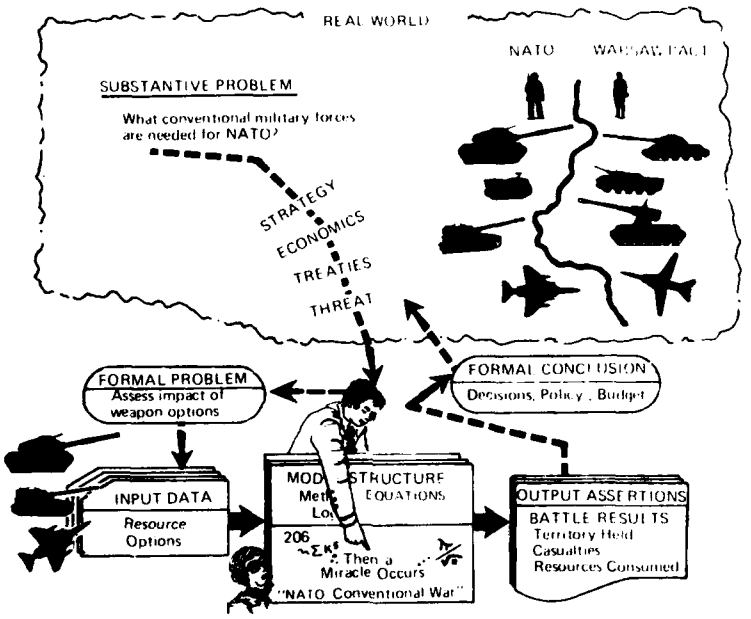
Squishy problems do not have a formulation that is both analytically tractable (based on science, empirical research, etc.), and which unambiguously captures the substantive problem (assures a good problem/model fit). Clear verbal statements of the problem may appear analytically intractable, while analytically tractable formulations of the problem are likely to significantly distort, or at least restrict, its substance.

Determining what conventional military forces will best serve U.S. needs in the NATO theater is an example of a very squishy problem. The substantive issue involves a host of political, economic, and military considerations--foreign policy, budget implications, and the Nation's war-fighting strategy, to name just a few. Some of the considerations--such as the qualitative-quantitative options for providing miltary capability--may be under the decisionmaker's control; others, such as the Warsaw Pact's military intentions, may not. In general, no single consideration predominates (military solutions yield to budget and foreign policy considerations, and vice versa), and no objective standards exist for comparing their relative importance. The formulation and conclusions for a squishy problem are inherently subjective--requiring and depending on careful and considered judgment by the decisionmaker.

Figure 2.3 shows the analysis of this problem using a policy assisting model that is a mathematical abstraction of a NATO/Warsw Pact conflict. As in figure 2.2, there is a substantive level at the top, and a mathematical level at the bottom. Unlike the rigorously quantifiable case, however, the relationship between the problem and the model may be highly tenuous and ambiguous. The model is rarely a direct match for the problem, nor the conclusion a straightforward translation of the results. As the figure attempts to illustrate, both the substantive problem and conclusion

may be far from sharply defined. The model itself will still be sharply defined. This is true even if the model is not completely specified (by science, empirical data, or the decisionmaker's judgment, etc.) but is left partially implicit (in which case the model builder's judgment must compensate for the missing information), since that is the nature of a mathematical model.

Figure 2.3
Analysis of a Squishy Problem



The links between the problem and model, and the corresponding links between the results and conclusion, may be far from clear. For this reason, it is useful to think of an intermediate formal level between the substantive and mathematical levels. The formal problem links the substantive problem and the model by specifying what parts of the problem are being modeled. It delineates the relationship of the model to the problem (the assumptions being made, the major categories left out, etc.). The formal conclusion serves similarly as a link between the output assertions and the substantive conclusions--interpreting the results in a context more closely tied to the substantive problem and less to the model.

If the substantive problem is rigorously quantifiable--as in the case of our artillery example (figure 2.2)--then the formal and substantive problems may merge together and assume a structure virtually identical to that of the model. As we move along the spectrum in the direction of increasing squishiness, however, the three become more clearly distinguishable: the model's assertions become more clearly insights, rather than conclusions, and the importance of the decisionmaker's judgment becomes increasingly apparent.

A potential for distortion

In squishy problems, the activities of formulation and interpretation are inherently judgmental--and inextricably linked. Formulation, from the substantive to the formal problem and from the formal problem to the model, is primarily a process of taking away--of removing pieces to make the problem smaller and more analytically tractable. Formulation may also involve some adding on, in the form of simplifying assumptions that are questionable on substantive grounds, but which make analysis easier. Interpretation, conversely, requires putting things back--adding in the omitted influences and compensating for any distortions that may have occurred. Interpretation requires a knowledge of the judgments embodied in the formulation, as well as an intuitive understanding of the substantive problem and the methodology being used.

To begin understanding the potential for distortion inherent in the formulation and interpretation of a squishy problem, it is useful to consider similar effects in the analysis of our NATO example. There are other terms that could be applied to the influences we are about to describe; however, for the present, it is appropriate to distinguish between these influences at the formal and mathematical levels of analysis.

Formal level -- context demarcation and interpretation

Figure 2.3 shows the formal level of analysis as a mental activity performed by the decisionmaker. The downward links, from the substantive to the formal problem, represent the process of assimilating the substantive problem, partitioning it, and specifying the context of the mathematical analysis. In the upward links, the decisionmaker interprets the results of the mathematical analysis, integrates the products of that interpretation with substantive considerations that were not included in the mathematical analysis, and then renders a formal conclusion. Observe also, that the figure's decisionmaker shows an interest in the model's

structure--an activity not clearly specified by the links. As we will see later, the decisionmaker's interest in the structure is well-founded and should be an important consideration at the formal level of analysis. For now, however, we want to concentrate on the links leading toward the decisionmaker--i.e., those links coming from the substantive problem and the model's output assertions. These are the generally recognized sources of information for decisionmaking.

It is fairly evident that the context of our model (battle dynamics) provides an incomplete fit for the substantive problem. Interpretation of the same results may vary between decisionmakers, and with the contextual demarcation of the substantive problem. The potential magnitude of these differences can be seen in our NATO example by considering the following questions (refer to figure 2.3).

- What defines victory--casualty levels, ground gained, or the control of strategic objectives? Over what time period--72 hours, a week, or ...?
- What is the purpose of NATO's conventional military force--to defend in a major war, or to defend in lesser conflicts that are below some predetermined nuclear threshold?
- Will France participate in a NATO conflict?
- What are the budget trade-offs between "on-line" and strategic reserve forces--does it cost more to station tank units in Europe, or to maintain them in the U.S. with an airlift capability? How will this affect the cost-effectiveness relationships between tanks and attack aircraft?

These questions make clear the potential for different conclusions about "What forces are needed for NATO?". They also point out several considerations that are fundamental to the analysis of a squishy problem. We elaborate only that

- the potential for distortion is due to an incomplete problem/model fit;
- the analysis depends to a great extent on the decisionmaker's skill and sensitivity in understanding real world processes; and
- the conclusions are the decisionmaker's conclusions, not the model's conclusions.

That the "decisionmaker-real world" link is critical to the analysis of our NATO example is without question. However, our discussion has thus far considered only the case of an incomplete fit between the context of the model and the substantive problem. A similar potential for distortion exists within the policy assisting model itself. Recall that these models are an admixture of science, empirical research, theory, and the builder's judgment--their mathematical structure is an incomplete fit for the context they purport to represent.

Mathematical level -- context formulation and interpretation

The model in our NATO example is a mathematical abstraction of theater-level war. It describes the progress of battle on a day-to-day basis, down to and including the ammunition expenditures and killer/victim relationships for each type of weapon in each maneuver battalion. But battle is a complex process, involving a staggering totality of considerations and intuitive judgments. It affects and is affected by variations in human behavior, command decisions, logistics, time-space relationships, and each unit's ability to shoot, move, and communicate. As a consequence, the model has to include simplifying assumptions and--where scientific knowledge is incomplete--mathematical judgments. Each of these assumptions and judgments is--in a sense--an incomplete problem/model fit, and each introduces a potential for distortion.

If this were a rigorously quantifiable (and validated) model of NATO war, the decisionmaker could accept the model's "battle results" as a real world fact; there would be no distortion in the mathematical analysis. But that is not the case.

> "Without adequate understanding of the empirical context, without full realization of the embedded assumptions, and without appreciation of exclusions, and omissions, a potential user is easily led down the garden path." 1/

An illustration of a few of the model's simpler assumptions serves to bring this point home.

1/See Brewer and Hall [13], p. 10.

- Exclusions and omissions. "Weather is not considered to affect activity on the ground and it is not an overt factor in aerial activities either. The obvious justification for ignoring the effects of weather is that they would apply to friendly and enemy forces alike." 1/ That is no doubt true, but it is also true that weather could markedly influence the nature and tempo of NATO/Warsaw Pact battle. (What if one side has air superiority? What about helicopter mobility of allied reserves, the air-ground reinforcement of defense sectors threatened by breakthrough, etc.?) "In some scenarios a slower advance by the attacker would permit the defender time to bring in additional troops which might well alter the outcome of the war." 2/

- Embedded assumptions. "Including elements in the model makes the implicit assumption that they are more relevant, more important, than those excluded." 3/ Inconsistent detail is especially deceptive.

  -- Shortcomings in the modeling of maneuver tend to bias the model towards favoring a weapon's firepower attributes, vis-a-vis its mobility and tactical employment characteristics. What is the value of the XM-1 tank's agility on the NATO battlefield?

  -- Ammunition shortages degrade unit performance. Ammunition consumption is measured instantaneously; resupply is modeled every 12 hours, and without a logistics network. How will this affect our results?

- Empirical context. "Since there is no generally accepted or independently validated 'theory' of war and combat, operational and campaign models have an ad hoc quality." 4/

1/See Dondero [25], p. E-17.

2/Ibid.

3/See Brewer and Hall [13], p. 38.

4/See Stockfisch [80], p.3.

Consider the attack of an enemy position. As the attack continues, both sides suffer casualties, and eventually one side discontinues or "breaks off" contact (rarely is the loser annihilated). In our model, this battle termination process is controlled by a set of user supplied breakpoints (the casualty percentages at which a side will withdraw or discontinue attack).

The choice of a breakpoint is very important. Not only does it affect who "wins" an engagement, it also has an important bearing on factors such as: battle duration, casualty levels, equipment losses, and ammunition expenditures. In fact, marginal changes in the value assigned to just one breakpoint may have the effect of drastically changing the policy interpretation of the entire model.

A second point, interlocking with the first, is that the model's breakpoint hypothesis (i.e., the assertion that casualty percentages are the cause of battle termination) is simply a mathematical convenience. The hypothesis itself " * * * yields theoretical implications that are at variance with the available battle termination data in several essential respects." 1/

But battles do end, and while science cannot explain why they end, the decisions on "What forces are needed for NATO?" must be made. Breakpoint choices and other human judgments are inherent in those decisions--with or without the model. The underlying imperative is to get the best possible NATO decision. The model--as an extension of judgment, and as an information-economizing device--has considerable potential to aid in that process. 2/

1/See Helmbold [47], p. v.

2/ This point has been made by G. H. Fisher in remarking, "This is especially true if sensitivity analyses have been made * * * [and] the final results are still within relatively narrow ranges. Given results of this kind, the decisionmaker can be less concerned about making a mistake regarding the quantitative aspects of the problems, and he may then feel somewhat more comfortable about focusing more of his attention on the qualitative * * * considerations." See Quade and Boucher [74], p. 40.

Yet, the model is not an oracle. Its results must be tempered by experience, intuition, and judgment. And here, we ask: "Are all interpretations equally credible?" Or, might that depend on

-- the decisionmaker's background and institutional affiliation;

-- his/her understanding of the model's structure and breakpoint assumptions; and

-- the source 1/ of those assumptions?

The message is simple and clear. Understanding the real world is one thing; understanding the model's representation of the real world may be quite another.

A policy assisting model's potential for distortion is not to be lightly dismissed. If one were to make even a casual survey of figure 2.3, the implications for decision-making at the formal level of analysis should be obvious. Credible conclusions depend not only on the "decisionmaker-real world" link, but also on the creation of a "decision-maker-model" link (the decisionmaker's understanding of the model's assumptions and structure). 2/

Recall Simon's admonition: mathematics is one thing; science, quite another. "A theory is defined by a set of assumptions concerning the relationships among a set of elements or variables * * *," 3/ and

- A policy assisting model is a form of theory--a quantitative structure of fact and conjecture,

---

1/Suppose, for example, the model's breakpoints are based on historical analysis. Does it matter "What war?" or "What battles?"; or how well the analysis was done?

2/We use the term "decisionmaker-model" link in contradistinction to a "decisionmaker-model output" link. The policy assisting model's mathematical structure is, by definition, an incomplete fit for the context it purports to represent. The model is an inadequate surrogate for the real world.

3/See Martin Shubik, in Brewer and Hall [13], p. 9.

fashioned to represent a problem that does not have a scientific or "objective" solution. 1/

Thus,

- The model's utility is not determined by truth criteria. The validation of a policy assisting model is tantamount to establishing a new scientific law. 2/ Because such models go beyond objective science, they cannot produce results that are "true" or "valid" in a rigorously quantifiable sense.

- The aim is to enhance and extend judgment. The model should be looked upon not as a "provider of solutions," but rather as a framework which permits science and judgment to be brought together and made explicit. It is the explicitness of this structure--the decisionmaker's ability to probe, modify, and examine "What if?" alternatives--that is of value in extending judgment.

- The conclusions are the decisionmaker's, not the model's. Policy assisting models have "considerable potential as an aid to judgments and a source of insight into and understanding about squishy problems. That potential is diminished significantly

---

1/By definition, the solutions provided by a policy assisting model are "suboptimal." Extreme care is required when translating model results into substantive conclusions.

2/There is little agreement even about primitive terms in the modeling community. The words validation, verification, appraisal, and evaluation are used interchangeably in practice although various efforts are being made to sort them out. Here, we use the terms "model appraisal" and "model evaluation" to denote activities undertaken to inform decisionmakers about the level of confidence that might be placed in a model's results. "Validation" and "verification" are important components of these activities; however, in this report we avoid the terms "model validation" and "model verification" as misleading to the extent that they divert scarce attention away from the fact that the policy assisting models are inadequate surrogates for the real world. This view is somewhat heretical and is probably not shared by those who would seek to confine a modeler's responsibility to the mathematical analysis--but then, the issue is more pragmatic than technical.

when the * * * [model] is looked to as a replacement for judgment and a source of objective knowledge." 1/

The rationale that says "decisionmakers don't have time to understand a policy assisting model" is a direct contradiction to the literal justification which argues that Government uses such models to compensate for the inadequacies of human judgment. 2/

The model is a theory, and a theory is no better than its assumptions.

"Unfortunately these assumptions, cast in mathematics and computer code, are usually embedded within a model and seldom explicated." 3/ "All

---

1/See Strauch [81], p. xiii.

2/The reader who is so inclined should have no trouble quibbling with our characterization of the decisionmaker as an expert and a direct user of policy assisting models. Indeed, there are *many* variations on this arrangement; both at the first level, and as information is filtered up through a decisionmaking hierarchy. Yet, regardless of the finer distinctions about how information flows to the decisionmaking arena (which in certain contexts are both subtle and important), the ultimate responsibility for determining its credibility rests on the decisionmaker. In this report, we have tried to keep things simple. Our characterization portrays the "in-principle" conditions which minimize the potential for distortion--a useful framework for discussion and understanding. The "in-practice" conditions are rarely this simple--nor is the potential for distortion. The diligent reader, therefore, might want to consider the implications of using an analyst as an intermediary between the decisionmaker and the model in figure 2.3.

--What is the influence of the analyst's skill and experience?
--Whose judgment is extended: the decisionmaker's, the analyst's, or some combination? How?
--Does the model guarantee an advantage over the traditional form of decisionmaking? If not, what does it take to make it so?

3/See Brewer and Hall [13], p. 9.

of the assumptions of a model must be made explicit. If they are not, this is a defect." 1/

A subjective solution

There is no clear-cut mathematical formula to describe a squishy problem. We simply do not have the scientific knowledge necessary to understand, or predict, its behavior. For processes which we have limited understanding, we formulate a hypothesis until further observation teaches us differently. Some term this "subjectivity" and denounce it soundly; others term it "wisdom" and recommend it highly. Whatever--this is the "quantified judgment" that gives us the power to seek "better" solutions to the really difficult public policy problems. But then, the question inevitably arises, "How do you know the hypothesis is sound?" In the area of reprocessing uranium and plutonium fuels, for example, how

> "* * * would you even begin to describe analytically the obviously important interrelationship between future reprocessing costs and the need to keep the risk of radiation contamination to workers and the environment to an acceptable level?" 2/

To be candid, science and mathematics don't know. The good sense of the decisionmaker generally provides the only guide.

The validity of the conclusions reached depends on: (1) the internal logical validity of the mathematical analysis, and (2) the empirical validity of the linkages between the model and the substantive problem. As is true with the rigorously quantifiable models, the logical validity of the mathematical analysis can be determined objectively, without reference to the builder's judgment. "This is not true, however, of the linkages between problem and model. In general, no objective standards exist by which these linkages can be validated." 3/

1/See Quade [73], p. 168.

2/The quote is Vince Taylor's [85], p. 51. His perspective on the role of mathematics, science, and the analyst in public policy formulation is worth reading.

3/See Strauch [81], p. 17.

To recapitulate, policy assisting models are

- not verifiable by an independent critic; and
- not "objective" in solving real world problems.

This fact--that we cannot guarantee the "real-world truthfulness" of a model's assertions--is extremely important. We trust a decisionmaker's judgment; but when that judgment is "extended" by a model--a model that uses unverified assumptions to go beyond science and "objective" fact--"How can the decisionmaker be sure that the model is, in fact, serving as an extension of his/her own judgment?" Where are the checks and balances, the safeguards?

Policy assisting models propose to intervene in the formulation of public policy. They demand most rigorous appraisal. It is in the common interest for decisionmakers to insist on assurances that:

- the model is mathematically correct,
- the part of the model which purports to be "objective" fact matches the real world, and
- the model uses empirically valid data.

And also, that a policy assisting model is transparent. 1/

TOWARD STRENGTHENING THE FOUNDATION FOR DECISION

At the outset of this chapter we asked whether the Foundation for Decision is mathematics, or science and expert judgment. Our reply is that it is all three. Their roles are complementary, yet separate and distinct.

The policy assisting model is an analogue of that foundation. And, to make perfectly obvious the two separate, yet

---

1/The user should be able to see and understand the model's logic, if not at a glance, at least with a limited amount of study. For policy assisting models, this includes that portion of the model's logic which goes beyond "objective" fact. The user should have available, for example, information regarding: the source of the model's assumptions, their real world interpretation, and the influence they exert on the model's results (sensitivity).

related issues that this report is attempting to deal with, let us point out that we have thus far focused on the decision-maker as a "consumer" of model information. Now, we want to consider the decisionmaker in an "investor/manager" role. We ask the very practical, "policy relevant" question: "What is being done to maintain or strengthen the Foundation for Decision?"

The investment of knowledge

The challenges confronting Government decisionmakers grow more awesome and complicated each year. Over the 1957-1977 period, the United States' real GNP increased at an average annual rate of 3.4 percent, doubling the Nation's total output of goods and services in less than 21 years. This means, figuratively speaking, that today's 21-year-olds are surrounded by twice as much of everything newly man-made as they were at birth. By the time they reach age 63, perhaps three such doublings will have occurred--and, since the increases are compounded, the Nation will be producing eight times as much as when they were born.

The benefits of this economic growth are well-known; less obvious are its demands, which:

- intensify demands on technology and finite resources;
- heighten society's complexity (as industrialization introduced concerns for education, retirement, and environmental issues); and
- increase society's vulnerability (to economic conditions, nuclear power generation, pesticides, etc.).

Each year, Government decisionmakers must assimilate more technological and sociological facts, evaluate more complex situations, and make more decisions, more rapidly than ever before. Today the problems are difficult: tomorrow they may be more difficult.

The aim is to make better decisions faster, easier, and cheaper. And here the idea is often advanced that the use of computers and quantitative methods provides an advantage over the traditional approach to decisionmaking, which relies on intuitive judgment. We support that idea.

In principle, a computer model

"and our degree of understanding of an associated environment can be enhanced by decomposing a context into subproblems, separate relationships,

and individual elements; by measuring important individual elements; by experimenting with a whole model to appraise and adjust it; and finally, by using the model for projective purposes." 1/

In the case of many of the problems encountered in Government decisionmaking, however, our knowledge is insufficient to get this far. And, as Bonder has ably explained:

> "Because of the absence of* * *[science and 'objective' knowledge], we in the modeling community have been developing models somewhat as natural philosophers or Platonists in that the models are developed by pure reasoning and logic alone. I strongly believe that models developed on this basis which are not experimentally verified cannot and should not be used as an evaluation mechanism to provide accurate, point estimate predictions * * * for use by decision makers. Rather, I think these intellectually developed models (rather than experimentally developed) should be used for analysis purposes to provide managers with
>
> (a) insights into directional trends to increase their understanding of the system dynamics, and
>
> (b) guidelines for the development of data collection plans (i.e., what data are important, how accurate must they be, etc.).
>
> "This kind of information is generated by parametric variation of the model variables and assumptions designed to answer 'what would happen if' questions and to expose the full range of possible effects of a decision." 2/

The difference between the modeling of rigorously quantifiable (the "in-principle" description) and Bonder's very squishy problems should be evident.

But we need not stop here. Suppose we think of the model itself as a beaker and the science and "objective" fact in its logic, as sand. The implications of this analogy are depicted in figure 2.4. On the left, the rigorously quantifiable model is shown as a beaker full of sand. As we

---

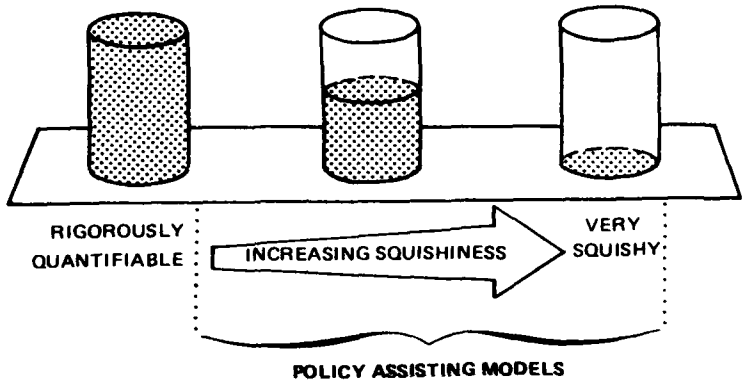
1/See Brewer and Hall [13], p. 5.

2/See Bonder [9], p. 77.

move to the right--in the direction of increasingly squishy problems--the policy assisting models are shown as containing proportionately less and less sand. Let us think of the transparent (not filled with sand) portion of a beaker as the model's assumptions.

Figure 2.4

An Analogy: The Model As a Beaker of Sand



We view the rigorously quantifiable model as "complete"; the beaker is filled with sand, and the model's logic is based on verified science and "objective" fact. In this instance, a decisionmaker would rarely be interestd in the model's logic. Nor would there be any impetus to improve the model's results. It could be relegated to the "shelf", and taken off and relied upon for "objective", factual solutions to problems as needed.

By contrast, we view a policy assisting model as "incomplete"; the beaker is only partially filled with sand, and the model's logic contains assumptions that go beyond science and "objective" fact. And let us say quite pointedly that, while a decisionmaker may not need to "see through the model's science," s/he has a vital interest in the transparency of its assumptions. This is the essence of the method. Succinctly,

- The model is a "mathematical blueprint"--not unlike an architect's plans for a new home. Each provides explicit details about the designer's choice of construction materials, and how they are organized

and related in a framework for decisionmaking. And each is a proposal, contingent upon the decisionmaker's acceptance or modification.

- In other words, the policy assisting model is a tool for integrating diverse forms of knowledge. It offers the opportunity to bring together the "best" of science and considered judgment.

  But without an adequate theoretical basis, who could reasonably expect all of the designer's assumptions to be correct? Whose judgment will be used to "build the house"?

- The point is that we have moved from the notion of a "sand-filled beaker" replacing human judgment, to an open, explicit analysis that relies on the decisionmaker's investment of knowledge.

  Yet, by definition, no judgment is known to be correct. The advantage of a "well-filled beaker"--and the impetus for improving the model--should be evident.

And, since the aim is to make better decisions faster, cheaper, and easier, let us also understand that

- The model's assumptions are testable hypotheses. They are mathematical assertions about the individual elements of a real world problem--and they should behave like the real world. Thus, when the real world is measurable, the model's assumptions can be tested and refined--sand can be added to the beaker. (An economic model's assumptions, for example, might be expected to produce the growth fluctuations, unemployment, and inflation observed in the real economy.) 1/

  But, if today's problems are difficult, and tomorrow's more difficult: "Will decisionmakers 'learn' from their experiences?"; "Will they elicit feedback, and compare performance with expectations?"; and finally, "Will they pass on the benefits of that knowledge to their successors?"

1/Obviously, these remarks pertain to "mathematical judgments." While it is unlikely that the "beaker" for any significant public policy problem will ever be completely filled with sand, each grain that is added removes one more element of uncertainty.

To relegate this model to the "shelf" is to imply there is no learning; to believe otherwise is a pitfall.

The point is this: the policy assisting model has considerable potential--not only as a tool for today's decisions, but also as a vehicle through which today's decisionmakers can strengthen the foundation for tomorrow's decisions. That potential depends significantly upon the decisionmaker's "investment of knowledge" and the management of quantitative tools.

The management of quantitative tools

Historical trends in the management of quantitative tools portend harmful effects; both for the method, and for important decisionmaking processes within the Federal Government. The critical problems, in our view, are not technological, they are institutional--and pretending they do not exist will not make them go away.

Garry Brewer has put a part of the problem in chronological order, and his remarks are telling:

> "In 1971, the GAO flatly stated, 'There are no Government-wide ADP documentation standards' [92].
>
> "In 1973, the GAO recommended that the Department of Defense * * * [establish] 'a requirement for periodic independent technical reviews of computer models to insure continued improvement in their development and employment as well as in the studies in which they are used' [93]. As far as can be determined, no such independent, periodic technical reviews have been undertaken.
>
> "In 1974, the GAO was still able to conclude, 'Our current study showed that the documentation guidelines at Federal agencies were still inadequate' [94].
>
> "In 1975, a National Science Foundation study determined that only about 20 percent of the nonmilitary models funded by the federal government could pass a minimal standard for documentation [34]." 1/

But we need not stop here.

1/See Brewer [12], p. 70.

- In 1976, we reviewed PIES-74 (an energy model used as the major policy analysis tool in the development of the 1974 Project Independence Report). We reported that the model was inadequately documented, and that it contained serious methodological problems warranting improvement. A single example will serve to illustrate:

  -- Any energy forecast has significant economic implications and should be consistent with economic forecasts. However, in PIES-74 the demand estimates were not integrated with the estimates of economic activity. This precluded a satisfactory analysis of the economic implications of an energy forecast. 1/

  It is also worth noting that: A congressional modeling specialist "discovered an FEA [Federal Energy Administration] mistake during a critical point in the debate by the 94th Congress on the implications of the President's decontrol program. FEA analysts had failed to recognize that a base case in running a model of the economy already included decontrol. The error resulted in a three-to-fourfold underestimation of the effects of decontrol on unemployment and real GNP." 2/

- In 1977, we evaluated TRIM (a model used throughout Government to estimate the dollar costs, caseloads, and income distributional effects of alternative welfare policies). Our findings included:

  -- There are some errors in the computer code which indicated inadequate verification of the computer model during its development.

1/Deficiencies in the management, use, and/or documentation of federally funded computer models have been cited in a number of our previous reports--including our evaluations of the 1974 Project Independence Evaluation System (an energy model), and the Transfer Income Model (a welfare model). See U.S. General Accounting Office [91]-[97].

2/See Greenberger and Richels [38], pp. 18-19. Notwithstanding any criticism, PIES-74 was a valuable first step toward providing an integrated framework for evaluating energy policies. Severe time constraints took their toll; the 1974 version of PIES is no longer used. The new Department of Energy has initiated actions that will provide for the regular review and evaluation of their models.

-- A number of versions of the model exist, and this increases the possibility that agencies using different versions of the model will make different estimates of the costs, impacts, and benefits of the same proposal.

-- In general, TRIM should not be used to provide absolute, or point, estimates; especially if no information is provided as to the uncertainty inherent in TRIM's estimates.

One Department took exception to this conclusion in replying, "We feel that the TRIM methodology incorporates current knowledge in these areas, and is the most precise method available for making such [point] estimates." 1/ That argument may have merit, but what about the other versions of TRIM that produce different answers? Why not the best?

The comments of one of the developers are also worth noting. "Most of the funding, even during the early development period, was for specific policy estimates. Because of insufficient funding by the government, we had a desperate struggle to complete the model with the effort on the brink of default. The story has been the same ever since, with many agencies wanting to use the model for policy estimates, but with money for maintenance, documentation and improvement very difficult to obtain. This problem is much broader than this model. * * * This process is exacerbated when a public good used by many agencies is the product. Everyone wants to use the public good for free and few feel a responsibility to support its continued development and maintenance." 2/

And the story goes on, affecting:

- Documentation. "Management did not clarify documentation requirements for the model. As a result, only

---

1/See U.S. General Accounting Office [97], p. 97.

2/See U.S. General Accounting Office [97], pp. 104-105. A TRIM User's Group has been established to deal with problems affecting the model's use and maintenance.

the developer understood how it worked and the relationships maintained by the variables incorporated into it." 1/

- Appraisal. "Many capabilities built into these devices have not been subjected to validation. Not only is their empirical base dubious or admitted to be lacking, but few efforts are being made to collect missing or questionable input data or to execute sensitivity analyses according to an appropriate experimental design. The lack of sensitivity analysis is related to deficiencies in estimating the validity of input parameters. Neither of these matters seems to be taken seriously." 2/

- Transparency. Modeling efforts have "changed in general from a relatively simple and 'visible' aid to judgment, to an esoteric and frequently unquestioned producer of * * * outcomes." 3/

- Analyst qualification. "The training issues are becoming increasingly important, at least as far as PIES is concerned. Right now it's not easy for a senior analyst to learn PIES in less than several months to the point where he (or she) can contribute to modeling or analysis. I'm talking about smart Ph.D.'s new to our staff; it's taking them on the order of five or six months to really understand what's going on." 4/

  There is a "strange inconsistency between people wanting more detail and yet resenting having to spend $120K a year to put three people to the task

---

1/See U.S. General Accounting Office [96], p. 8. The National Bureau of Standards has recently published Guidelines for Documentation of Computer Programs and Automated Data Systems, Federal Information Processing Standards Publication 38. While these are merely guidelines, we believe they represent a reasonable approach to computer documentation.

2/See Shubik and Brewer [77], p. 62.

3/See Honig [50], p. I-4.

4/See Gass [35], p. 120.

of understanding the model and how to use it. At the same time, we've got billion dollar investments riding on decisions, and if the models produce a 10% piece of information for inputs in decision making that's good and well worth the money." 1/

- Data. "The input data used in models often have an obscure or unknown empirical foundation, and the relevance of much data (even when it is valid) to the [problem] * * * is unknown." 2/

By contrast, today there are healthy trends toward improving the management of quantitative tools. 3/ The effort deserves constant oversight--for it is the decisionmakers who create the institutional pressures for improvement.

The aim and opportunity

When attention is directed to the needs of Government decisionmakers, the opportunity to apply fruitfully the techniques of quantitative methodology would seem to be much greater than ever before. Demand for the product is certainly high, and its future should well be a productive and beneficial one.

But models have to be kept in context. (For an interesting description of how models are used by the Joint Chiefs of Staff, see appendix I.) In formulating public policy, only human judgment can cope with the complexities of modern socioeconomic behavior; the appropriate role for quantitative methods is to enhance and extend a decisionmaker's judgment. This is their aim and opportunity.

In turn, this requires that the decisionmaker be

1. a knowledgeable consumer of model information:

   - Quantitative decisionmaking is advantageous--only when it embodies, rather than replaces, expert judgment and "objective" fact.

1/See Office of Naval Research [69], p. 38.

2/See Stockfisch [80], p. 11.

3/See Gass [35]; Office of Naval Research [69]; U.S. General Accounting Office [101]. See also the following footnotes in this report: fn 2, p. 26; fn 2, p. 27; and fn 1 ,p. 28.

- Application of the method can be made in the name and rhetoric of science, but without being subjected to the normal evaluative standards of science.

- Profound mismatches exist between the methodological tools, and the current quality and quantity of theory and supportive data.

2. an investor of knowledge:

   - In the analysis of rigorously quantifiable problems, the quantitative method uses science and "objective" fact in lieu of human judgment.

     --The objective is to predict real world behavior.

     --The focus is on results derived from the model.

   - In the analysis of squishy problems, the quantitative method uses human judgment in lieu of science and "objective" fact.

     --The objective is to extend judgment.

     --The focus is on the open explicit structure for analysis.

3. a concerned manager:

   - Regardless of the finer distinctions about how information flows to the decisionmaking arena, the ultimate responsibility for determining its credibility rests on the decisionmaker. It is in the common interest for decisionmakers to oversee the modeling effort (see figure 2.5), and to insist on assurances that:

     --the model is mathematically correct,

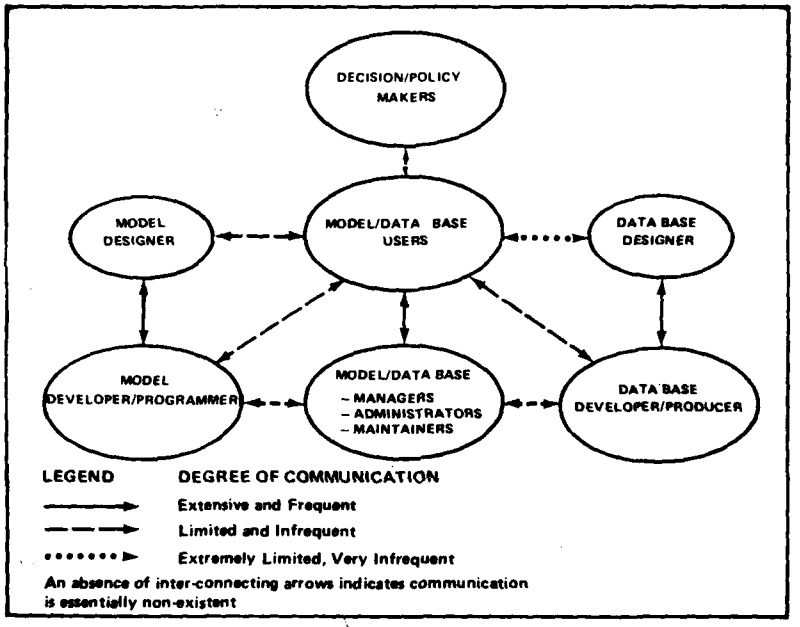
     --the part of the model which purports to be "objective" fact matches the real world, and

     --the model uses empirically valid data.

     And also, that a policy assisting model is transparent.

Figure 2.5

Relative Degree of Communication Between Decisionmakers and Users, Designers, Developers, Producers and Managers of Models and Data Bases



SOURCE: OFFICE OF NAVAL RESEARCH, *THEATER-LEVEL GAMING AND ANALYSIS WORKSHOP FOR FORCE PLANNING, VOLUME I-PROCEEDINGS*, "OPENING REMARKS, SESSION I" 1979.

These facts are inescapable. If we look to the results of computer models for

> "* * * answers to complex social, political, or behavioral problems * * * we may be returning full circle--back to the shaman and the oracle, asking for a magical mechanism beyond the range of human consciousness and understanding." 1/

1/See Strauch [81], p. 87.

CHAPTER 3

THE MODERN DESIGN FOR DEFENSE DECISION

This chapter presents a highly simplified description of the philosophy and analytical approach underlying the planning, programming, and budgeting of U.S. ground and tactical air forces. Its basic purpose is to provide an overview of, and a context for, the individual questions of theory and technique discussed in subsequent chapters.

PERSPECTIVE

Since the early 1960s, the United States has evolved "a new philosophy, technique, and style of defense management. To some extent this was inevitable; military planning today presents a new problem, different from earlier military planning, not in any deep logical or philosophical sense, but in a practical sense. The radical change in weapons, with their almost exponential increase in complexity, and the concomitant need for research and development, forced a new emphasis on science and engineering and rendered past military experience a far less certain guide to future conflict. Central to this new concept of defense management is the acceptance by decisionmakers of policy advice provided by systematic analytic studies. Such studies * * * have thus become an essential part of the policymaking process." 1/

The management philosophy

The foundation for this "Modern Design for Defense Decision" was established during Robert S. McNamara's tenure as Secretary of Defense. His remarks of that period attest to its philosophy.

> "I consider the budget nothing more and nothing less than the quantitative expression of a plan or a policy. So in developing the budget I propose to start with the plan or the policy and translate it into quantitative terms, terms of benefit and cost.
>
> "The creation of the Department of Defense resulted from the clear recognition that separate land, sea, and air warfare is gone forever. * * * Our international political problems and our military problems are now indivisible.

1/See Quade and Boucher [74], p. 1.

"The President's charge to me was a two-pronged one--to determine what forces were required and to procure and support them as economically as possible.

"Our problems of choice among alternatives in strategy and in weapon systems have been complicated enormously by the bewildering array of entirely workable alternative courses which our technology can support. * * * The difficult question is 'What is required?'

"We first took a major step forward in the development of our planning, programming, and budget process.

"To be really meaningful the defense program must be looked at in its entirety with each of its elements considered in light of the total program. This can only be done at the Department of Defense level. For example, the size of the POLARIS force cannot be determined in terms of the Navy shipbuilding program or even the entire Navy program, but can be validly judged only in relation to all the other elements of the Strategic Retaliatory Forces--the B-52s, the ATLAS, the TITAN, and the MINUTEMAN ICBM's. Similarly, the requirement for Air Force tactical fighters cannot be determined independently of the requirement for Army ground forces.

"To make such a review a reality, a 5-year program was devised presenting the proposed force structure and cost projections in terms of the principal missions of the Defense Department.

"In our approach we show just what we are planning to spend on each mission, such as for the strategic retaliatory forces, * * * general purpose forces * * * etc. These categories are further broken down into individual systems and projects. For each mission, you can see how many planes we plan to have, how much investment is involved, what the expected operating costs are, how many personnel are involved. In each case, competing programs and systems are judged on the basis of their contribution to the mission to be accomplished and to the Defense effort as a whole.

"The judgment inherent in this balancing of programs and systems can no longer be intuitive or rely on past experience alone. The range of choice is too broad; the number and type of alternatives too great.

"In the selection of weapon systems, in the design of forces, and in determination of the level of the national defense effort, therefore, we are making greater use of a technique called systems analysis.

"These are two of the primary management tools we put to work--a mission-oriented planning and programming process to assist in defining and balancing the total effort, and systems analysis to assist in the selection of specific weapons systems and courses of action from among potential alternatives. But management tools and techniques are only that--they assist, but only assist, in the decision-making process." 1/

In what follows, we describe these tools.

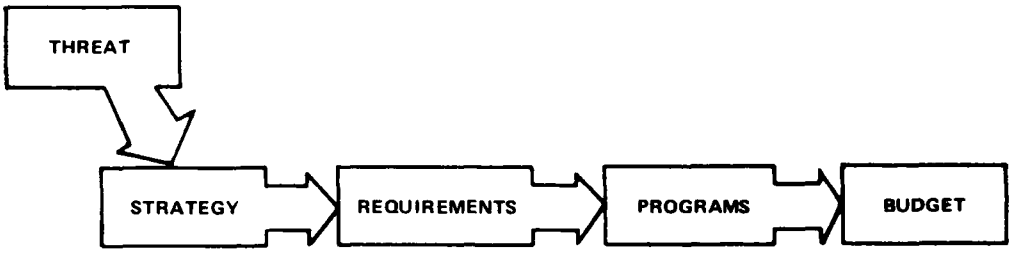
The planning-programming-budgeting system

The Planning-Programming-Budgeting System (PPBS) can be summarized in a few words. The strategy is developed in consideration of the threat. Force requirements are developed to support the strategy. Programs are developed to provide, in an orderly and economical fashion, the manpower and weapons needed over a period of time. And finally, the budget is developed in such a manner as to get the most defense out of any given level of available resources.

1/See Robert S. McNamara, in Tucker [87], pp. 9-15.

Figure 3.1

Planning, Programming, and Budgeting



1. Planning. The first phase of the PPBS sets the pattern for the entire process. Planning starts with the assessment of the threat to the security of the United States and culminates with the projection of force objectives to assure the security of the United States.

   - Joint Chiefs of Staff (JCS) submit their strategy to the Secretary of Defense (SECDEF).
   - SECDEF issues strategic guidance.
   - JCS submit their forces plan to SECDEF based on the strategic guidance. This plan is not fiscally constrained but presents what is needed and what can be attained.

   As seen, the major portion of the planning effort is accomplished within the JCS area. The concept is to assess the world situation (friend and foe) at prescribed future time periods, technical capabilities required, military strategy to counter threats to the national security, and to state force objectives to satisfy the national strategy.

2. Programming. The programming phase translates the approved concepts and objectives, prepared during the planning phase, into a definitive structure expressed in terms of time-phased resource requirements including men, monies, and materiel. This is accomplished through systematic approval procedures that "cost out" force objectives for financial and manpower resources 5 years into the future, while at the same time displaying forces for an

additional 3 years. This gives the Secretary of Defense, the Congress and the President an idea of the impact that present-day decisions have on the future defense posture.

- SECDEF issues fiscal guidance to the Services and DOD Agencies for each of the 5 program years.

- JCS submit joint force recommendations (Joint Force Memorandum) with rationale and risk assessments. These are fiscally constrained consistent with SECDEF fiscal guidance.

- Services and DOD Agencies submit their program objectives (Program Objective Memorandum, or POM) to SECDEF including forces and support, with rationale and risk assessment. These are also fiscally constrained consistent with SECDEF fiscal guidance.

- SECDEF issues final program decisions after draft decisions have been commented on by the Services and DOD Agencies.

  --Assistant Secretary of Defense (Program Analysis and Evaluation) prepares Issue Papers for SECDEF which analyze the Service POM's and evaluate the costs and capabilities of alternative programs.

  --Issue Papers are circulated to the JCS and DOD Agencies for comments before being submitted to SECDEF for decision.

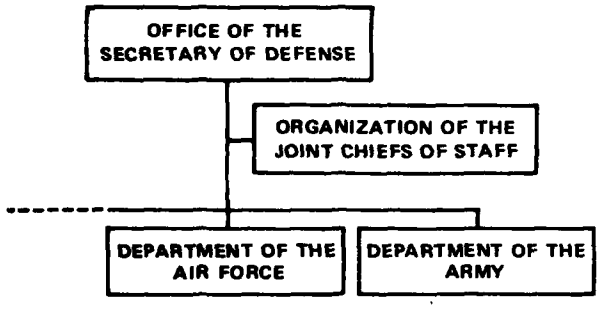
3. Budgeting. This is the final phase in the Planning-Programming-Budgeting system. The annual budget expresses the financial requirements necessary to support the approved forces and programs set forth under the first program year of the Five Year Defense Program.

Implicit in the process outlined so briefly are the development of mid-range objectives, the conduct of special studies, and research and development of weapons systems and their support. In fact, all the resources of the Services are drawn upon to formulate their plans, programs, and budgets.

We pause here to focus explicit attention on three aspects of the PPBS that bear directly on subsequent discussions in this report. With the reminder that we are talking about conventional ground and tactical air forces, we will address these points in the context of the principal analytical groups shown in figure 3.2.

Figure 3.2

Principal Analytical Groups Supporting POM Development and Review



ANALYTICAL GROUP
Assistant Secretary of Defense (Program Analysis and Evaluation)
Studies, Analysis, and Gaming Agency
U.S. Army Concepts Analysis Agency
U.S. Air Force Studies and Analysis

- First, note that each of these analytical groups uses a different policy assisting model to analyze the force requirements for theater-level war.
- Second, note that the analytical groups supporting SECDEF and the JCS have the principal responsibility for analyzing the distribution of resources between Services.
- Finally, clearly recognize that:
  - -- In the planning phase, the JCS's force plan is intended to provide a required military capability at the least cost. (Effectiveness is fixed; cost is variable.)
  - -- In the programming phase, SECDEF provides fiscal guidance. The JCS's Joint Force Memorandum and the Service POM's are intended to provide the maximum military capability for a fixed cost. (Effectiveness is variable; cost is fixed.)

As we shall see on pp. 39-42, there are substantial conceptual differences between these two approaches.

Still, there is one additional topic to be considered in our discussion of the PPBS--the Services' acquisition of major weapon systems. On this matter, the Services manage the acquisition process on a daily basis, while the Secretary of Defense, apart from the PPBS cycle, oversees the development/acquisition of individual weapons through a series of milestone reviews. Decisions to initiate, modify, or terminate a development/acquisition program are made by the Secretary. Those decisions are incorporated in the Service's POM and then enter the normal PPBS cycle. For the purposes of this report, we are interested in the relationship between the performance of these new weapons and the military capability implied in the POM. We elaborate:

- Analyses performed in support of the acquisition process tend to focus on the engineered performance characteristics of competing weapons as opposed to their contribution to combined-arms battle. The analytical groups and models for these studies differ from those shown in figure 3.2. The point is that the criteria used for selection in a lower level problem should be consistent with the criteria used at higher levels. There are conceptual differences between the "effectiveness" criteria used in weapon system acquisition and the POM.

- In 1978, we reported that the identification of the need for a new weapon system was a potential problem. We recommended that the Secretary of Defense:

  -- "Require that each service justify each new planned, major weapon system by showing how it relates to a recognized deficiency in a mission area." 1/

THE ANALYTICAL ENGINE

The second management tool "put to work" under Secretary McNamara was systems analysis.

1/See U.S. General Accounting Office [99], p. 26.

The central problems in the design of analyses to aid decisionmakers involve selecting operationally meaningful objectives, measures of their attainment, and criteria. In what follows, we attempt to provide a brief understanding of these relationships--and of the difficulties in definition and measurement which are introduced in going from simple decision problems concerning narrowly defined systems and operations to complex decision problems involving broader issues. Portions of our remarks are based on L. D. Attaway's "Criteria and the Measurement of Effectiveness." 1/

Cost-effectiveness analysis

It is an established practice in DOD to support decisionmaking with systematic analyses of alternatives in terms of their military worth and cost. All of these cost-effectiveness analyses embody the following major elements: 2/

OBJECTIVE:	What we desire to achieve
ALTERNATIVES:	Competitive means for achieving the goal
COSTS:	Expenditures to acquire each alternative
EFFECTIVENESS SCALE:	Scale indicating degree of achievement of goal
EFFECTIVENESS:	Position on effectiveness scale assigned to each alternative (by measurement)
CRITERION:	Statement about cost and effectiveness which determines choice

The effectiveness scale provides the "yardstick" by which the various alternatives are compared; effectiveness is the ability of an alternative to achieve an objective in terms of that scale. This effectiveness scale is crucial--the benefits of the same alternative will differ according to the "yardstick" used to measure it.

In some cases the selection between alternatives is easy. An extreme case of this--shown in figure 3.3--occurs when one alternative is more effective at every cost. Even though "dominance" designates Alternative II as preferred, the required level of effectiveness must be specified before the preferred level of investment can be selected. 3/

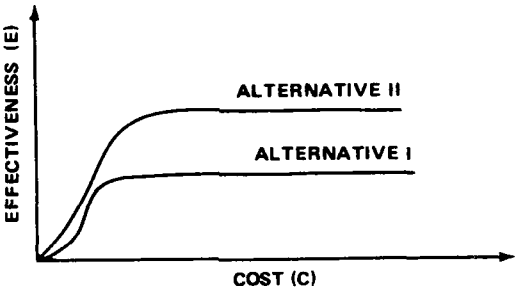
1/For an excellent discussion of this topic: see Attaway, in Quade and Boucher [74], pp. 54-80.

2/Ibid., p. 55.

3/See Attaway, in Quade and Boucher [74], pp. 58-59.

Figure 3.3

Dominance

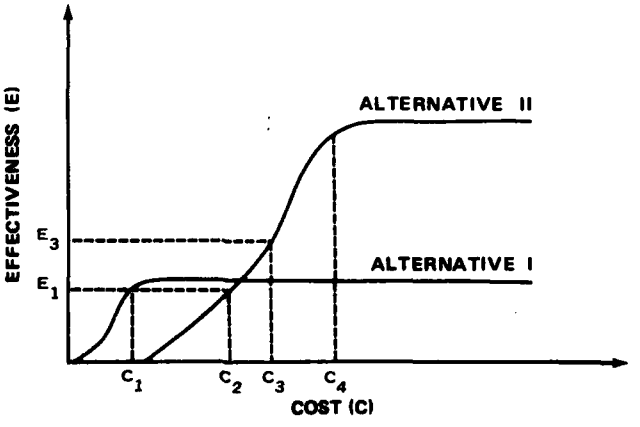


Source: L.D. Attaway in Quade and Boucher[74].

But rarely is the choice that simple. Let us consider a simplified decision problem--the selection of a new tank gun. The objective is to increase the tank's antitank capability. Two alternatives are available: a 90mm gun (Alternative I) and a 120mm gun (Alternative II). The effectiveness scale is the range at which the gun has 90 percent probability of scoring a first round kill. For each alternative, the cost scale is the cumulative expenditure (on propellant, warhead, and gun technology) for achieving that level of effectiveness. Figure 3.4 illustrates the cost-effectiveness curves for this idealized example.

Figure 3.4

Cost and Effectiveness



Source: L.D. Attaway in Quade and Boucher[74].

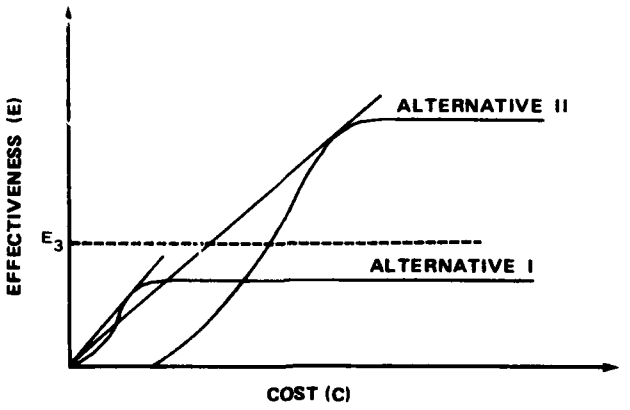
Notice that a clear choice between the two guns is not possible without more information. What is missing is some knowledge of why the improved antitank capability is needed. Although Alternative I achieves only a modest level of effectiveness (E_1), it does so at about one-third of the cost of Alternative II. If the level E_1 is adequate, why not select Alternative I and thereby minimize cost? 1/

Cost is often limited to some level such as C_2, in which case Alternative I is the obvious choice. On the other hand, if the goal is to achieve some minimal new level of effectiveness, such as E_3, no matter what the cost--then Alternative II is the obvious choice. A related point is that if C_3 is a reasonable cost to pay, then the case for C_4 is stronger since substantial gains in effectiveness can be made for a relatively small additional investment. 2/

Somewhere in the middle are criteria that specify neither required cost nor effectiveness. Maximizing the ratio of effectiveness to cost seems to be a workable criterion, for example, since increasing effectiveness (while decreasing cost) is intuitively appealing. Nevertheless, as figure 3.5 illustrates, this criterion has a serious defect. 3/

Figure 3.5

Effectiveness/Cost Ratio



Source: L.D. Attaway in Quade and Boucher [74].

1/See Attaway, in Quade and Boucher [74], p. 56.

2/Ibid., pp. 56-57.

3/Op cit., p. 57.

Since the effectiveness-cost ratio for either alternative is simply the slope of a line drawn from the origin to a given point on the curve for that alternative, and since the ratio is greatest at the "knee" of the curve, Alternative I is clearly preferred with this criterion. However, if E_3 is the minimum level of acceptable effectiveness for the guns, then Alternative II is the obvious choice. The point to be made here is that such criteria suppress absolute levels of effectiveness and cost--considerations which may be very important to decisionmaking. 1/

In general, it is not possible to choose between alternatives just on the basis of cost and effectiveness data. Usually, either a required effectiveness must be specified and then the cost minimized for that effectiveness; or a required cost must be specified and the effectiveness maximized--but not both. One cannot maximize effectiveness and minimize cost simultaneously. In figure 3.5, for example, minimum cost corresponds to zero effectiveness, and maximum effectiveness corresponds to a very large cost. 2/

The point is that the PPBS is not a paragon of economic theory. The JCS's force plan, the Service POM's and the cost and operational analyses of new weapons alternate between the objectives of minimizing cost and maximizing effectiveness. This is not to say that the PPBS has no value, but merely to draw explicit attention to the difference between the economic theory and the practicalities of real world management. In our opinion, both the PPBS and cost-effectiveness analyses are useful management tools.

> "The practical application of these tools to the problems of decision in the Defense Department remains much more a matter of judgment and common sense and much less of esoteric techniques that most of the published literature would suggest. The value of the development of these tools has been not in their rigid mechanical application but in providing a new way of looking at problems." 3/

The question remains, "How do we measure the effectiveness of a weapon system?"

1/See Attaway, in Quade and Boucher [74], p. 57.

2/Ibid.

3/See Alain C. Enthoven, in Tucker [87], p. 7.

A hierarchy of analysis

> "In the environment of the Department of Defense, the kinds of problems to which we try to apply cost-effectiveness analysis are largely these: force, composition, R&D, the selection of weapons, and the development of preferred manpower and logistical policies. These are different in a very marked degree from those operational problems studied in World War II. There is less emphasis today on the best tactical employment of weapons and much more upon the major decisions of how to allocate resources to various force mixes, [and] to the development and procurement of one weapon as opposed to another." 1/

In World War II, operations research activities were focused on improving the combat effectiveness of existing weapons. The availability of operational statistics contributed to the success of that work in two important ways. First, it made the problem largely one of statistical inference; and second, it provided the analytical setting needed to evaluate the weapon's effectiveness in the context of real world battle. In other words, operations research dealt with rigorously, or at least reasonably, quantifiable problems.

Today, the situation is more complex. There are no combat statistics for new weapons, but numerous decisions must be made during their development. Figure 3.6 illustrates a hierarchy of analysis supporting this process.

1/See Grosse [39], p. 3.

Figure 3.6

A Hierarchy of Weapon System Analysis

Level of analysis	Assessments of effectiveness	Trade-offs
Battle		
↕ Units Weather Topography	Casualties Movement Rates Weapons Attrition	Numbers and Types of Unit Unit Tactics
Encounter		
↕ Multiple Weapons Types Ceiling/Visibility Terrain	Targets Killed per Day/Sortie Exchange Ratios	Weapons Mixes Sub-unit Tactics
Engagement		
↕ Single Weapons Types Controlled Conditions	Probability of Killing a Single Target Probability of Loss to Single Attritor Type	Weapons Tactics and Design Features
Operations		
↕ Weapon Properties	Endurance Range/Speed Fire Rate	Design Specifications
Engineering		
Weapon Configuration	Weight Power Size	Design Features

Source: Extracted from *Indices of Effectiveness in General Purpose Force Analysis*, The BDM Corporation, 1974.

Notice how the assessment of a weapon's effectiveness changes between the "engineering" and "battle" levels of analysis. At the "engineering" and "operations" levels of analysis, the problems are, in general, reasonably quantifiable. As we cross the line into and through the "engagement" level, however, the problems become increasingly squishy. 1/ (The term systems analysis is often used to connote the application of operations research methods to squishy problems.) Since no combat statistics exist for a new or proposed weapon, its performance in battle is often simulated through

1/Our concern here is with the "effectiveness" side of a cost-effectiveness relationship. Cost is also a squishy problem. The alert reader will recognize that the choice between investment cost, operating cost, 5-year program cost, and life-cycle cost (including any discount rates) can have an important influence on the results of analysis.

the use of a combat model (a policy assisting model). 1/ But care must be taken--going from the simpler problem involving "weapon configuration," to the problem of "weapon effectiveness" in combined-arms battle is difficult; and inversions in the analytical results can occur. For example:

- At the engagement level, two armored personnel carriers (APC's) are compared--one has a telescoping mount to fire a TOW antitank rocket; and the other has an elevated, but fixed, mount. Both weapons are designed to use the cover of an embankment for hull protection while firing on enemy tanks. In the engagement, the telescoping mount, even though it has a higher unit cost, would be more cost-effective because it lowers the weapon's silhouette (thereby minimizing detection and hits by enemy tanks).

- At the encounter level, the "better" APC (with the telescoping mount) and friendly tanks fought against an enemy tank force. But the telescoping mount jammed in the "up" position after firing. Because the "jammed" TOW mount was exposed, the enemy considered it a high threat weapon, forcing them to maneuver and split their fire. The friendly tanks--subjected to less hostile fire--inflicted a higher loss ratio on the enemy. In this case, the effectiveness of the friendly force increased. Should we have paid the higher cost for the telescoping mount, or is there a flaw in the analysis?

1/DOD uses a number of combat and campaign models to simulate phases of conventional war, from small engagements to major campaigns. In a 1971 survey, for example, Shubik and Brewer identified over 450 active models, simulations, and games within DOD. In 1973, we reported that the costs of building a representative sample of 104 DOD models totaled $28.8 million--the cost of individual models ranged from about $1,200 to $3 million. But determining the full costs of this modeling effort is deceptively difficult. The work involves many invisible costs, overhead costs, jointly shared facilities, and jointly used products; and formulating a meaningful costing procedure poses deep problems that are far from being resolved. In this report we focus on the application of quantitative methodology. We simply point out that DOD's modeling requires a considerable investment of time and resources. For a discussion of these costs, see: Shubik and Brewer [77]; and U.S. General Accounting Office [93].

Figure 3.7

Use of the CEM Model in Defense Decision

USE	PURPOSE	GIVENS	OUTPUT	IMPACT
JSOP (Joint Strategic Objectives Plan)	Influence National defense planning	1987 Threat National strategy	Force requirements Objective force	Shows minimum goals acceptable to JCS Shapes interservice balance
TA (Total Army Study)	Supports Program Objectives *Memorandum* (POM)	1984 Other programs Threat Combat force Consolidated Guidance	Program force	Sizes Army
OM (Omnibus Study)	Capability review	1978 Other budgets Threat Real world force	Current force deficiencies	Quick fixes Adjustments in early program years
Ammo Rates (Ammunition Requirements Study)	Determines ammo war reserves Distributes ammo assets	1984 Threat Combat force Weapon capabilities	Expected Expenditure of Ammo (EEA), by type	Sizes Army reserves Scales production base
WARF (Wartime Attrition Replacement Factors)	Determines equipment war reserves	1984 Threat Combat force Weapon capabilities	Expected equipment losses, by type	Sizes port Authorizes Acquisition Objective (AAO)
WARRAMP (Wartime requirement for Ammunition, Materiel and Personnel)	Both above, plus personnel replacements	1984 Threat Combat force Weapon capabilities	Both above, plus personnel losses	Both above, Personnel, relacement policy
IDOFOR (Improved Definition of Objective Force)	Illuminates force design options	1990 Estimated resource constraints	Alternative 1990 force structure	Inferences of current threat

Source: U.S. Army Concepts Analysis Agency

From this point of view, the cost-effectiveness relationships at the lower levels of weapon system analysis are examples of what is called lower-level "suboptimization" in systems analysis. "It is the seeking of efficiency in the small, while total force structure analysis aims at efficiency in the large." 1/

The point is that the costs, complexities, and synergistic effects of modern weapons are requiring DOD decisionmakers to take a broader look at resource allocation issues. And, because many of the crucial issues must be resolved in the context of modern battle (featuring the combined effects of infantry, artillery, air, and armor), there is a vast increase in the numbers of variables to be considered. Defense planners are becoming increasingly dependent on some information-organizing device--such as the theater-level combat model.

Theater-level combat models provide Defense decisionmakers with insights about resource requirements (the personnel, ammunition, and equipment needed), and the results of a large force confrontation, usually in a specific geographic setting, such as the NATO Central Front. Figure 3.7 illustrates the extent to which one such model--CEM (CONAF Evaluation Model)--supports Defense Decision.

It is important to note, however, that no single model or analysis drives an individual decision. Rather, groups of studies tend cumulatively to foster a consensus on major issues. In doing so, the analytical efforts interact strongly with external realities and decisionmakers' opinions. The result of this interactive process is a gradual, continuing, and apparently strong influence on the resolution of major issues.

In what follows, we describe the four principal models of conventional air and ground warfare that are used in DOD force planning--and in the development and review of the Army's POM (see figure 3.2). These are

LULEJIAN ..... used by the Assistant Secretary of Defense (Program Analysis and Evaluation);

1/See G. H. Fisher, in Quade [73], p. 269.

IDAGAM II	used by the Studies, Analysis, and Gaming Agency (SAGA) of the Organization of the Joint Chiefs of Staff;
VECTOR-2	to be used by SAGA; 1/ and
CEM IV	used by the U.S. Army Concepts Analysis Agency.

1/VECTOR-2's data base is under development.

CHAPTER 4

HARD-NOSED DEMANDS FOR MODEL TRANSPARENCY, APPRAISAL, AND CONSISTENCY CAN STRENGTHEN THE FOUNDATION FOR DEFENSE DECISION

The Department of Defense employs a number of computer models to assist in the analysis of issues related to the planning, programming, and budgeting of U.S. conventional forces. This chapter presents a simplified overview of four theater-level combat models (policy assisting models) involved in that process. Our purpose is twofold:

1. to illustrate the diverse assumptions employed in the modeling of attrition--a principal combat process which drives casualty levels and resupply/ reinforcement requirements; and

2. to emphasize the need for ensuring that policy assisting models used in Defense Decision are

   - Transparent so that a decisionmaker can understand and use the model as an extension of his/her own judgment. Implying that

     --Assumptions are clearly described and held to manageable proportions, and

     --The deductive process leading to the model's assertions is clear (transparent).

   - Appraised so that a decisionmaker can be assured that

     --The model is mathematically correct,

     --The part of the model that is science matches the real world, and

     --The model uses empirically valid data.

   - Consistent so that communication is facilitated throughout the decisionmaking hierarchy. Implying that

     --Problems are analyzed in the same context, and

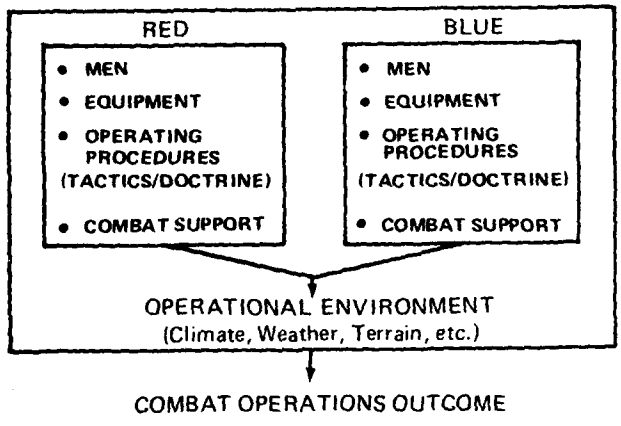
     --Differing viewpoints can be discussed on the basis of specific assumptions.

THEATER-LEVEL COMBAT MODELS-- A CAPSULE DESCRIPTION

To begin understanding what theater-level combat models do, it is useful to review some basic concepts (see figure 4.1). 1/

Figure 4.1

A Concept of Combat



Simply stated, all combat involves the interaction between two opposing forces, designated RED and BLUE. The forces are composed of men and equipment, are governed by operating procedures, and involve some measure of combat support. Both forces function in an operational environment, which is composed of natural factors, such as weather and terrain. The interaction between RED and BLUE both affects, and is affected by, factors such as:

1/Portions of our discussion are taken from Larry Low's "Concept and Plan" for the Theater-Level Gaming and Analysis Workshop, sponsored by the Office of Naval Research, 27-29 September, 1977. For an excellent description of the breadth of theater-level gaming, see Low [62]. For a discussion of the state-of-the-art, see Office of Naval Research [69].

COMMAND AND CONTROL

- Mission
- Composition of force
- Supporting units (artillery, aviation)
- Battle plan
- Time of battle
- Posture (hasty defense, etc.)

LOGISTICS

- Resupply
- Transportation
- Medical
- Equipment repair
- Construction (roads, bridges, etc.)

ATTRITION (COMBAT LOSSES)

- Weapon/Target characteristics (air-to-air, air-to-ground, ground-to-air, ground-to-ground)
- Munition characteristics (e.g., precision-guided or free-flight, fragmenting or solid shot)
- Engagement characteristics (visibility, movement, range, terrain, etc.)

The interaction between RED and BLUE results in a combat operations outcome, which is routinely measured in a variety of ways:

Movement of Forward Edge of Battle Area (FEBA)

- Territory controlled

Attrition

- Personnel
- Weapon (destruction and damage)

Resource Consumption

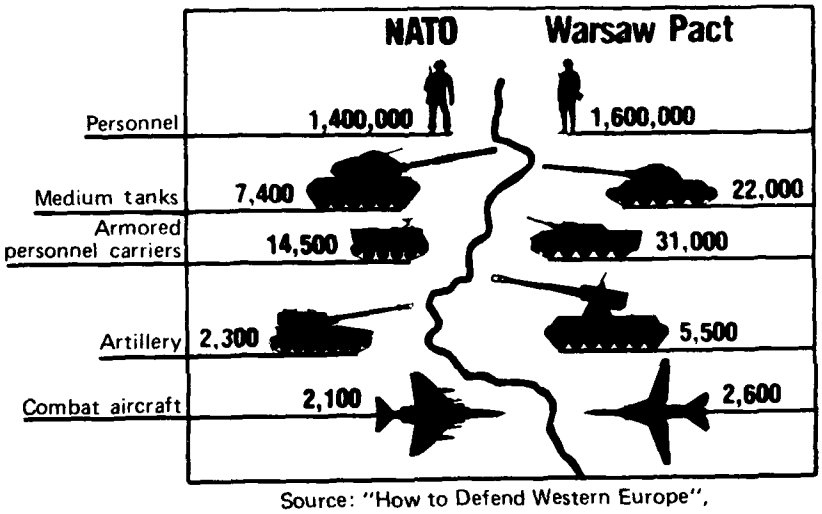
- Ammo expended
- Fuel used
- Supplies consumed
- Equipment lost

Models cited in this report

The models discussed in this report--CEM IV, IDAGAM II, LULEJIAN, and VECTOR-2 1/ --are analytic models 2/ of theater-level combat involving major confrontations between forces composed of diverse combat elements in a specific geographic area such as the NATO Central Front (see figure 4.2).

Figure 4.2

NATO/Warsaw Pact Force Comparison



Source: "How to Defend Western Europe", Fortune, Oct. 9, 1978.

1/For the remainder of this report, the models will be referred to simply as CEM, IDAGAM, LULEJIAN, and VECTOR. For model documentation: see CEM [15], IDAGAM [51], LULEJIAN [63], and VECTOR [102].

2/DOD uses many different types of combat models for analysis. Analytic models represent war via mathematical formulas. Simulations act out discrete combat processes. Both simulations and analytical models operate without human intervention. Human participation is included with interactive models and manual or computer-assisted war games.

VECTOR-2 is an analytic hybrid model which combines the properties of an analytic model with those of a simulation. The model acts out in detail combat processes such as the movement of front-line units, while other processes such as attrition are determined by mathematical formulas.

To represent an entire theater in the detail necessary to make useful statements to assist decisionmaking, the models require large amounts of input data (see table 4.1). 1/

TABLE 4.1

COMPARATIVE INPUT REQUIREMENTS

Combat Model	Number of Input Data Items Required for Comparable Runs
IDAGAM	40,000
VECTOR	212,250

Source: Command and Control Technical Center

Model outputs often take the form of periodic status reports 2/ that give a time history of battle evolution. Iteration of this process over a specified number of days produces a time history of war. Such models might be used to provide insights into questions concerning:

- systems choice (choice between comparable weapon systems),
- systems mix (choice between noncomparable weapon systems),
- force structure (choices in organization or force mix), and/or
- force level (how much is enough?).

1/Model inputs usually express technical and engineering attributes of equipment and munitions, behavioral characteristics of people who operate the equipment, terrain, etc. Note that RED force data represents a substantial investment in foreign intelligence. The difference between IDAGAM and VECTOR input data items is not indicative of source data requirements. VECTOR is a detailed model using essentially raw data. IDAGAM's inputs are highly aggregated.

2/Status report period varies between models--variations range from 1 to 24 hours, or more.

A general comparison of aggregated and detailed combat models

Aggregated and detailed models embody fundamentally different representations of combat. CEM, IDAGAM, and LULEJIAN are aggregated models; whereas VECTOR is a "detailed" model.

Aggregated models "lump together," or aggregate, similar types of weapons into a composite index, 1/ which is then used to represent the combat power of a military force (its ability to inflict and sustain casualties).

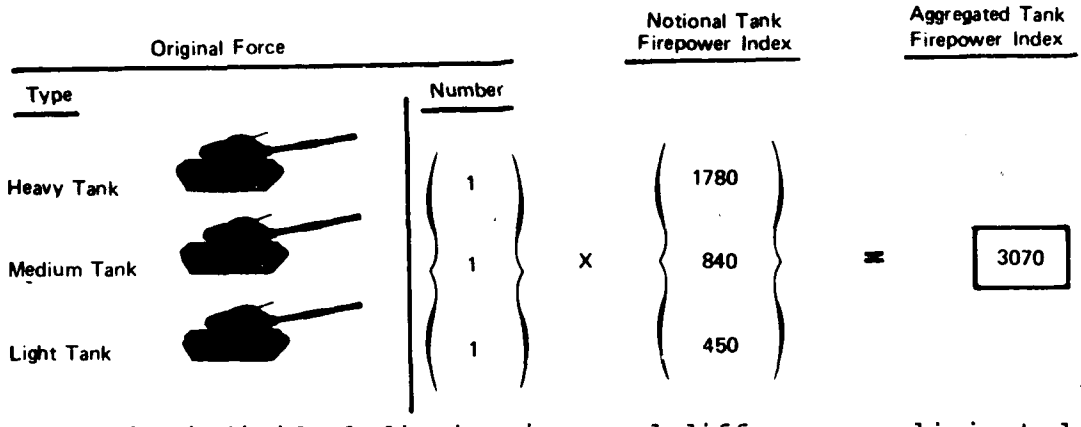
- The aggregation process differs between models. 2/ For example, CEM aggregates into three scores which represent a unit's ability to damage a particular type of target (e.g., personnel, light armor, and armor).
- The ratio of the composite index for BLUE to the index for RED--termed the force ratio--is the dominant factor in computing attrition rates and FEBA movement.
- Figure 4.3 illustrates this "lumping together." Understand, however, that aggregative measures mask distinctions. Aggregated models cannot reflect the subtle differences between similar weapons in a mixed force --e.g., between tanks such as the XM-1 and M60A3, or antitank weapons such as the TOW and DRAGON.

1/Firepower scores are commonly used as a basis for aggregation. The basic problem in developing an aggregation scheme is a linear weighting problem (e.g., how many rifles are equivalent to a tank, a flamethrower, or an aircraft?). Further, the linear addition of firepower scores does not reflect the generally accepted principle that the whole of a force is worth more than the sum of its parts--e.g., two tanks operating in unison should be more effective than if they were employed on independent missions. For a critical review of Firepower Potential Methodology: see Stockfisch [79], and Bode [7].

2/For further discussion: see Dondero [25], Honig [50], and Walker [107].

Figure 4.3

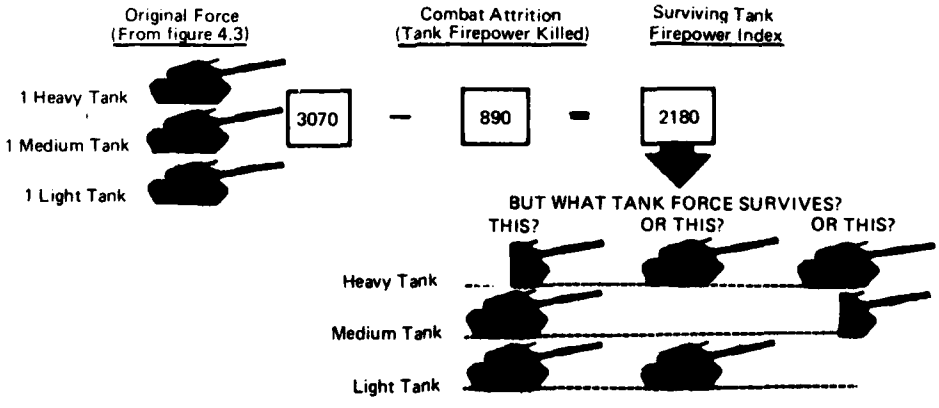
Illustrative Tank Aggregation



- The individual distinctions and differences eliminated by aggregation cannot be uniquely recaptured, or disaggregated. For example, when the firepower index is reduced by combat attrition, "what weapons survive?" and "what do we need to replace?" (see figure 4.4).

Figure 4.4

Illustrative Tank Aggregation/Disaggregation



Thus, with the use of aggregated models, the determination of "victims" and conversely, their "killers" is inextricably tied to judgment.

In contrast, detailed models present a more definitive representation of combat. The outcomes of maneuver force

engagements and the precise attribution of killer/victim relationships are predicted on: 1/

- Weapon performance factors such as firing rate, projectile flight time, probability of a hit, probability of a kill given a hit, etc.
- Acquisition parameters for visual and pinpoint acquisition of targets.
- Line of sight data for current terrain.
- Number of participating weapons of each type.
- Data describing force employment (initial deployments, rules for mounting and dismounting APCs, movement rates, open fire ranges, etc.).

These factors also embody judgments; but in the context of decisionmaking, they are very different than those embodied in an aggregated model.

The point is that detailed models make judgment explicit (and hopefully transparent). The decisionmaker can control target engagement priorities, open fire ranges, etc. Critical parameters can be modified to reflect changes in tactics and battle doctrine.

This is not so for aggregated models. The firepower index--keystone of the aggregated models--is predicated on a highly stylized interpretation of combat. Its derivation rests on judgments about tactics, open fire ranges, rates of fire, and the distribution of that fire (e.g., the percentage of tank firings directed at armor, mechanized personnel carriers, and foot infantry), etc. 2/ In other words, both the structure (aggregation scheme) and input data (firepower potential scores) for an aggregated model contain critical assumptions--assumptions that may be methodologically and intuitively inappropriate for a particular analysis.

Recall our concern for model transparency.

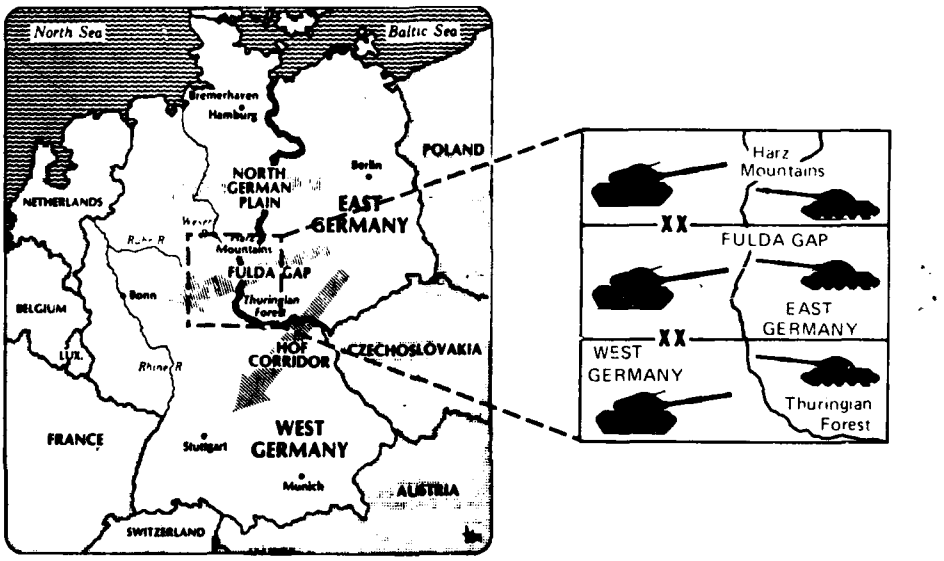
1/See Vector Research, Incorporated [103].

2/For example, if BLUE's firepower index were developed to reflect the distribution of fire against a tank-heavy RED force, the same value might significantly distort the results of an encounter with an infantry-heavy RED force.

The role of the combat sector

To model a NATO confrontation, the theater is divided into smaller geographic/tactical subdivisions--sectors. (Three such sectors are illustrated in figure 4.5).

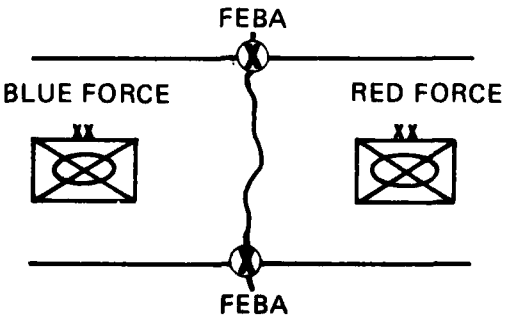
Figure 4.5
Three Combat Sectors of the NATO Central Front



A sector is designed to contain a specific force, often a corps or division, as illustrated in figure 4.6.

Figure 4.6

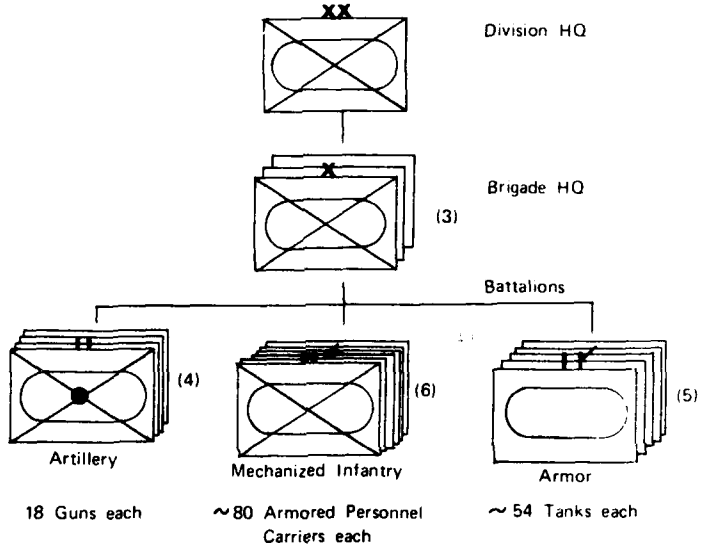
A Mechanized Infantry Division Sector



Two opposing forces (not necessarily the same size) face each other across the FEBA. The larger force contained in a sector--e.g., a division with roughly 17,000 personnel--is composed of numerous tactical subunits (see figure 4.7).

Figure 4.7

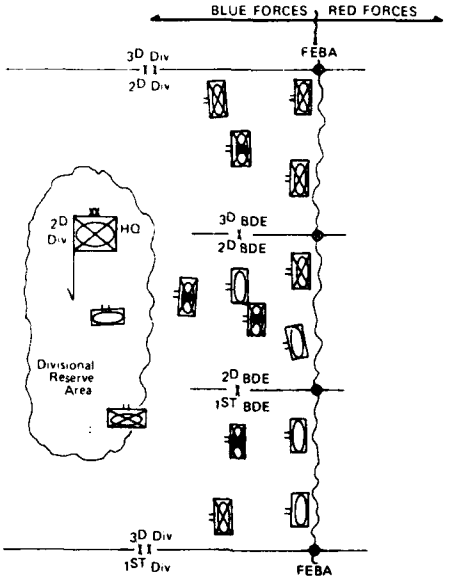
A Notional Mechanized Infantry Division



In our illustration (see figure 4.8), the Division is deployed with "3-Brigades Up." Each brigade, in turn, occupies about one-third of the Division's frontage and extends rearward to the depth necessary to harbor its two "on-line battalions," its brigade reserve, and its artillery battalion. The Divisional reserve and Headquarters are located to the rear of the brigade areas.

Figure 4.8

Notional Deployment of Blue Forces



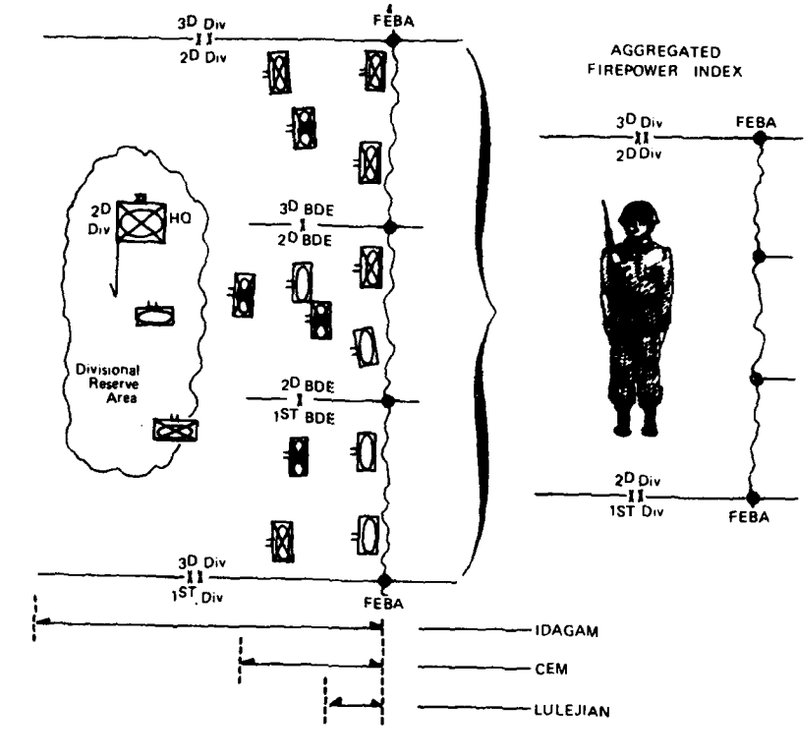
The role of sectors is nontrivial. As surrogates for more detailed positional information, sectors determine "interaction eligibility." That is, sectors provide inputs to mathematical models on the numbers of resources participating in combat interactions that are not, in general, further divisible. The combat results are determined by sector, and then summed over all the sectors to obtain theater-level results.

In VECTOR, the deployment of figure 4.8 is represented down to the specific weapon types in each maneuver battalion, with reserves, artillery, and supporting aircraft modeled to user specifications. The aggregated models represent this deployment with an index number that is based on the eligible occupants of the sector--as illustrated by the soldier in figure 4.9. 1/

1/Recall, the combat power of a force is represented by a composite firepower score which may be subdivided to denote the ability to destroy particular types of targets.

Figure 4.9

Portion of Notional Forces Included in Aggregation Schemes



The point is that aggregated models are insensitive to the spatial distribution of combat elements within a sector (the index number effectively depicts the force as if it were massed and deployed uniformly along the FEBA).

Even so, the "organization for battle" varies between models--the interpretation of who fights where is very different. CEM aggregates forces at the brigade level and resolves conflict across that brigade's frontage. IDAGAM and LULEJIAN both resolve conflict across the division's frontage; but IDAGAM aggregates the total division's force, while

LULEJIAN aggregates only the units adjacent to the FEBA. 1/ In short, each of the models uses different mathematical criteria to fight the same real world battle. Can we expect them to produce consistent results?

Battle is fought in both time and space, but we have thus far addressed only the spatial representation of forces. Now let us describe how the models represent time (see table 4.2). The process is basically one of breaking time into discrete intervals or periods--the "initial conditions" at the beginning of a period are used to predict what will happen during the period, and the predicted results become the "initial conditions" for the subsequent period.

Table 4.2

Time Periods Represented in the Models

	VECTOR	IDAGAM	LULEJIAN	CEM
Number separate time intervals	8	1	1	4
Smallest possible time period	30sec.	24hr.	24hr.	12hr.

Combat results are evaluated in the smallest possible time period. Thus, IDAGAM provides a "daily" calculation of battle results; while VECTOR simulates the battle in detail, updating the status and location of forces on a 30 second basis. Both VECTOR and CEM employ additional time intervals to model such factors as command and control, communications, and resupply/reinforcement.

The importance of time-space relationships to the real world NATO defense is unequivocal. In the event of battle,

---

1/Karr ([54], p.12) reports that IDAGAM's "developers envision a ground force unit as being of division or brigade size (although other choices are possible * * *)." The fact is that the model has not been completely documented. Available documentation indicates that the model's time intervals, spatial relationships, and casualty curves were designed to represent a division-size force. If this is so, the use of this model for other force sizes would be methodologically inappropriate and intuitively without justification.

NATO's forces would be deployed spatially (refer to figure 4.8)--an in-depth defense designed to gain time and economize forces. If Pact forces massed for a breakthrough, NATO's front-line units would have to "hold out" until reinforced by highly mobile reserves (note the special importance of attack aircraft, armed helicopters, tanks, etc.). This time to reinforce is critical.

Here we see another difference between the models' representation of battle. In the aggregated models, reserves must be committed at the beginning of battle, or they must wait until the next "day" (12 or 24 hours, depending on the model). VECTOR, on the other hand, will commit reserves according to user specifications. How will these time-space differences affect the results of an analysis? Will the decisionmaker's conclusion be independent of the model used?

Recall our concern for model consistency.

Attrition modeling

The approach used to model attrition dominates all facets of a combat model. Not only does it determine "winners and losers" and the movement of the front-line, it also drives casualty levels, equipment losses, ammunition expenditures, and resupply/reinforcement requirements.

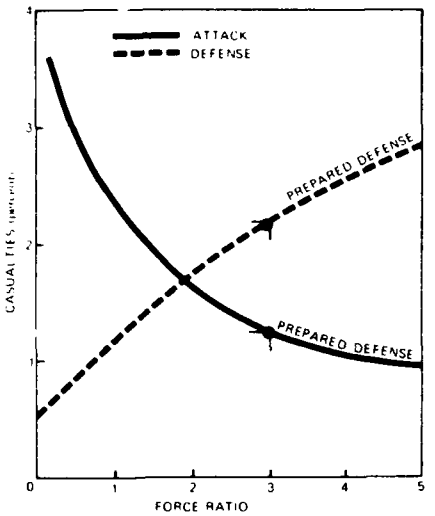
Index-number approach

Recall that aggregated models represent the "combat capability" of a force by an index number, and that combat causes attrition of these numbers. In IDAGAM, this attrition is determined through the use of casualty-rate curves, similar to those shown in figure 4.10. The curves relate casualty rates to force ratios and other tactical factors (e.g., attack or defense, and posture). 1/

1/IDAGAM is technically capable of representing six postures. Only four are described in model documentation--normal attack-delay posture, attack of a defensive position, breakthrough of a defensive position, and attack through a minefield.

Figure 4.10

Casualty-Rate Curves of the Type Used in IDAGAM



Note: To simplify the graph, curves for other posture/mission combinations are omitted. Casualty rates are in daily percentages.
Source: Research Analysis Corporation, NATO Combat Capabilities Study, Vol. IV, "Impact of Losses and Replacements on Unit Combat Capability."

To illustrate the use of these curves, let us consider the attack of a prepared defense. If BLUE attacked with a 3:1 force ratio over RED, then BLUE would suffer 1.3 percent personnel casualties; and RED (the defender), 2.1 percent. These are daily Division percentages--recall IDAGAM considers time in 24 hour periods and aggregates forces at the Division-level. Weapon-system losses--albeit much mathematical ado--are assessed as a pro-rata share of the Division's personnel casualties.

Of the four models, IDAGAM is the only one that uses historical data in the direct assessment of attrition. The model's casualty-rate curves are based, in a rather tenuous fashion, on historical data from 37 engagements in World War

II and Korea. 1/ But can historical curves represent the modern battlefield?

- The Army's "Active Defense" is a relatively new tactical innovation that is explicitly designed to change the shape of attrition curves. In theory, our defensive forces would sustain fewer casualties and inflict heavier losses on the enemy at all force ratios. Should historical curves be used to model this defense?

- Is battle a symmetric process? Will Warsaw Pact forces also use the "Active Defense?" Will the casualty curves be the same, irrespective of which side is the attacker?

- How are the weapon-system-performance characteristics of modern weaponry--e.g., precision guided munitions--related to the aggregated-force casualty rates of prior wars?

The other aggregated models--LULEJIAN and CEM--calculate attrition by means of mathematical formulae.

LULEJIAN calculates directly the attrition of tanks, armored personnel carriers, and personnel. Losses of other weapons (e.g., antitank weapons) are assessed in proportion to one of these categories. The basic premise of the model is that separation distances and losses are functions of one another. That such trade-offs occur in combat is plausible, but the extent to which LULEJIAN represents them appears questionable. 2/ The attrition equations are based on theory; their justification (or derivation) is not provided in model documentation.

1/More precisely, the data base for the four sets of curves is inordinately small and questionably representative of Division-level battles on the European landmass--11 engagements in Korea, and 26 in World War II (2 in France, 9 in Italy, and 15 on Okinawa).

2/See Karr [55], p.15. "To summarize, we believe that the assumptions inherent in the attrition/FEBA computations appearing in the Lulejian-I model are so incompatible with physical reality that they render the model significantly and unpredictably inaccurate as a representation of combat."

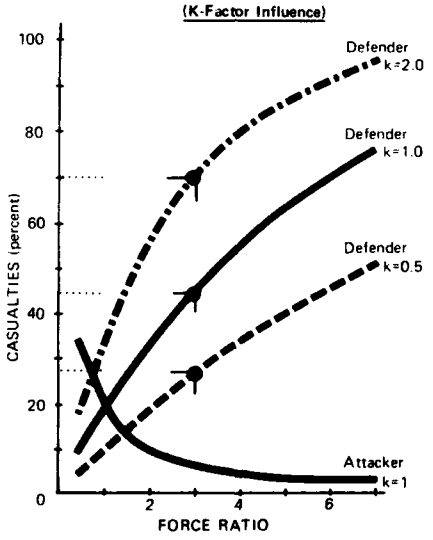
CEM is also a "weapon-killer"--with the primary victims being tanks, armored personnel carriers, helicopters, artillery tubes, and dismounted infantry. Casualties for these weapon-systems are determined independently by an exponential attrition equation that considers three factors: (1) the shooters' firepower index; (2) the number of targets (potential victims); and (3) a calibration parameter (the k-factor).

CEM's attrition equations are based on an admixture of science, theory, and a great deal of professional judgment. The firepower indices and the number of targets give a basic shape to the model's casualty-rate curves--but the k-factor dominates attrition levels. To bring this point home, we illustrate the k-factor's influence on the defender's casualty-rate curves for a hypothetical tank-versus-tank battle (identical tanks) in figure 4.11. (Not shown is a similar, but less pronounced, influence which would tend to drive the attacker's curve downward.) Consider the difference in the defender's casualty levels when attacked at a 3:1 force ratio.

Figure 4.11

CEM Casualty-Rate Curves

(K-Factor Influence)



The k-factor is a beneficial attribute. It is used to calibrate CEM's attrition levels against the results of a war game, field experiment, or high resolution model. In other words, attrition in this model can be adapted to serve as an extension of the decisionmaker's judgment. Of the four models, only CEM offers this as a "built-in" design feature.

Understand, however, that the k-factor is not derived from measurable physical parameters. It is obtained by solving the attrition equations in reverse--finding k-factor values that match the results of a specific engagement. 1/ Thus, it embodies--as assumptions--all of the weapon relationships, tactics, etc., that were present in that engagement. The point is that marginal changes in the k-factor can drastically change the policy interpretation of the entire model. The values used for the k-factor, and the scenarios from which they were derived, are not documented.

To summarize: IDAGAM's casualty-rate curves are based on a tenuous extrapolation of history; LULEJIAN's representation of combat appears questionable; and CEM's attrition levels are specified by an undocumented k-factor. Recall that these models are intended to serve as an extension of the decisionmaker's judgment--but how can they? "All of the assumptions of a model must be made explicit. If they are not, this is a defect." 2/

Lanchester-type approach

Lanchester's theory of combat is another approach to attrition modeling. This theory, when first introduced by Frederick W. Lanchester in 1914, was an attempt to describe the effects of concentration in warfare by means of a set of differential equations. The equations have come to be known as:

- Lanchester's linear law, representing combat where there is no concentration of force (area fire; shooters do not know when a target is killed); and
- Lanchester's square law, representing the effect of concentration (aimed fire; shooters know when a target is killed, and concentrate fire on the survivors).

To illustrate the effects of these laws, let us assume that "One man employing a machine-gun can punish a target to the same extent in a given time as sixteen riflemen. What is the number of men armed with the machine gun necessary to replace a battalion a thousand strong in the field? Taking

---

1/Model documentation briefly recounts k-factor methodology: see Louer [61], Part II, pp. 146-150.

2/See Quade [73], p. 168.

the fighting value of the rifleman as unity, let n = the number required." 1/

- Under linear law conditions, "as when searching an area or ridge at long range, or volley firing at a position, or 'into the brown',* * * the value of the individual machine-gun operator becomes * * *that of the sixteen riflemen that the power of his weapon represents." The number of machine-gunners required is

$$n = \frac{1000}{16} = 62.5$$

- Under square law conditions, "the enemy will concentrate on the one machine-gun operator the fire that would otherwise be distributed over four riflemen, and so on an average he will only last for one quarter the time, and at sixteen times the efficiency during his short life he will only be able to do the work of four riflemen in lieu of sixteen, as one might easily have supposed." The number of machine-gunners required is

$$n = \sqrt{\frac{(1000)^2}{16}} = 250$$

To consider the effects of concentration on cost-effectiveness analyses, we elaborate the square law's state-equation as

$$R_o^2 - R^2 = E(B_o^2 - B^2),$$

where E is the relative effectiveness of a BLUE troop to the relative effectiveness of a RED troop; and where B_o and R_o are the initial number of BLUE and RED troops respectively. The point is that

> "the effective strength of one side is proportional to the first power of its efficiency and proportional to the square of the number of combatants entering the engagement. Two opposing forces are then equally matched when the exchange rate is equal to the square

---

1/The example is Lanchester's. See Newman [68], pp. 2146-2147.

of the ratio of the number of combatants. Consequently, it is more profitable to increase the number of participants in an engagement than it is to increase (by the same amount) the exchange rate (by increasing the effectiveness of the individual weapons). This is not an argument against increased weapon efficiency; it is simply a statement that a tactical or strategical use of concentration may counterbalance any moderate advantage in weapon efficiency." 1/

A comparison of the quantitative-qualitative relationships in square and linear law battle will serve to bring this point home. Consider, for example, a duel-to-annihilation between 60 Warsaw Pact tanks and 20 NATO tanks. In figure 4.12 we illustrate the results of this battle, contrasting NATO tank effectiveness against the surviving percentage of the Warsaw Pact force. Without concentration of fire (linear law)--

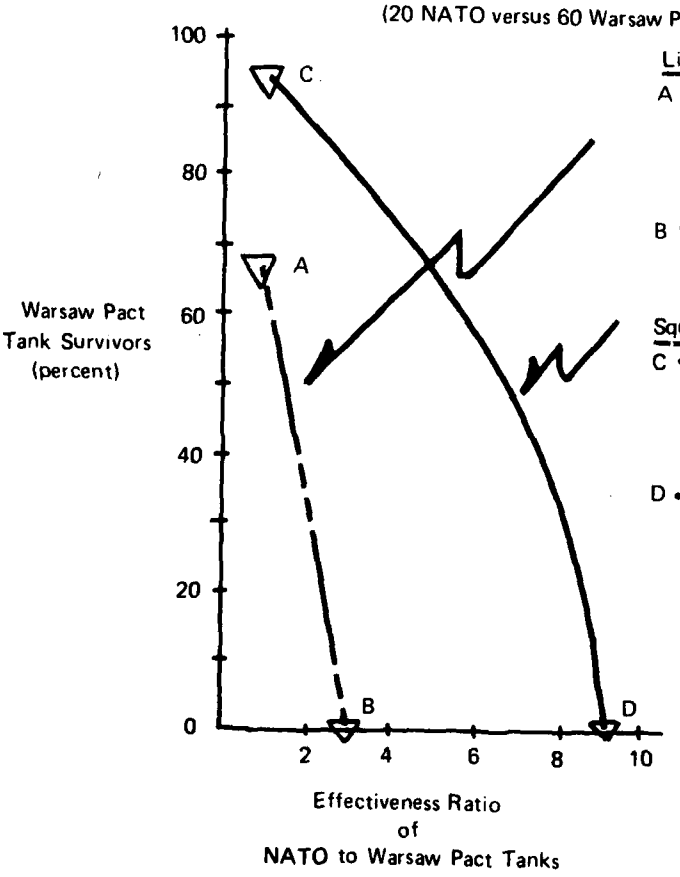
- If NATO and Warsaw Pact tanks are equally effective (point A), the NATO force is annihilated and 67 percent of the Pact force survives.
- If the NATO tank is three times as effective as the Pact tank (point B), the forces "break even"--they are both annihilated at the same time.
- If the NATO tank is more than three times as effective (not shown), the battle outcome reverses--the Warsaw Pact force is annihilated, and some portion of the NATO force survives.

But with concentration of fire (square law), the Pact's numerical superiority presents a more serious threat. Now, the NATO tank must be nine times as effective for the forces to break even (point D); and more than nine times as effective for the NATO force to win.

1/See Morse and Kimball [66], p. 65.

Figure 4.12

Quantity Versus Quality
in Linear and Square Law Attrition
(20 NATO versus 60 Warsaw Pact tanks)



Linear Law:

A • When NATO and Warsaw Pact tanks are equally effective, the NATO force is annihilated with 67% of the Warsaw Pact tanks surviving.

B • When the NATO tank is three times more effective, both sides are annihilated at the same time.

Square Law:

C • When NATO and Warsaw Pact tanks are equally effective, the NATO force is annihilated with 95% of the Warsaw Pact tanks surviving.

D • When the NATO tank is nine times more effective, both sides are annihilated at the same time.

Observe the clear difference between square and linear law attrition. Under the linear law, a weapon's "worth" is directly proportional to its effectiveness; under the square law this is not so. The choice of attrition profoundly affects the results of cost-effectiveness analyses--and consequently Defense Decision.

You would expect, therefore, that the models would agree on the basic form of attrition for a specific real world process. If that is what you expect, you are in for a grievous disappointment. In CEM, for example, surface-to-air missiles attrite attack aircraft by a linear law equation; LULEJIAN uses an exponential approximation of the square law for the same process. Which is correct? To be candid, we do not know. Quantitative techniques, including, especially, rigorous field experimentation have not been systematically applied to achieve these discoveries.

Lanchester's original work described a greatly simplified war--a force carried only one type of weapon, and the effects of terrain, tactics, supply, reinforcement, etc. were all ignored. The theory has been enriched since, and is very complex today. Lanchester-type equations--products of this enriched theory--have been extended to include combat between heterogeneous forces and such operational factors as:

- target acquisition considerations,
- range-dependent weapon-system capabilities,
- suppression and other temporal variations in fire effectiveness,
- the effects of logistics constraints,
- unit deterioration due to attrition, and
- unit breakpoints.

The VECTOR model effectively embodies the state-of-the-art in Lanchester attrition theory. "There is greater fidelity to the physical reality of combat, and also fewer arbitrary aggregations, fewer arbitrarily imposed sequences of events, and more reliance on definable and measurable inputs." 1/ Attrition is modeled explicitly and dynamically.

1/See Karr [56], p. 72.

A prominent feature of VECTOR is that the model uses basic data about the target acquisition process, range, intervisibility, projectile flight time, probability of a kill, etc. to "engineer" the attrition process. Now, obviously, this approach produces a highly complex model with an enormous appetite for data. That is true of VECTOR. But it is also true that the approach uses data that can be measured in laboratories and field experiments. And, to the extent this is done--to the extent that the data is validated as a pseudo-physical constant:

1. it does not require updating (as do most forms of aggregated data), and

2. it is eliminated as a decisionmaking variable.

In other words, VECTOR's attrition process--presently criticized for its complexity--has the potential, with time and supervision, of becoming reasonably transparent. 1/

Epilogue

The credibility of a theater-level combat model depends upon one's acceptance of the model's assumptions, and their source. It is in the common interest that a decisionmaker knows both.

This capsule description has focused on a single combat function: a force's ability to shoot, or more elegantly, to bring fire upon the enemy. We have shown how four models render profoundly different interpretations of the same real world phenomena--thereby underscoring the necessity that all the assumptions of a model must be made explicit. Note that our simplified illustration has excluded many of the assumptions used in "bringing fire upon the enemy," including the effects of:

- Battlefield obscuration--the reduction of visibility due to the presence of smoke, explosive debris, aerosols, etc. How will this affect attrition? Will all weapons, say tanks and attack helicopters, be equally affected?

1/By transparent, we mean that the users of the model, and the decisionmakers employing its derivative products, can see and understand those parts of the attrition logic which have not been empirically validated--if not at a glance, at least with a limited amount of study.

- Synergism--"The machine gun is a very lethal weapon, so lethal that infantrymen cannot ordinarily stand up to it. The effect of the machine gun, therefore, is not to cause heavy casualties, but instead to drive the infantry to the ground where they can become victims of the less lethal but more pervasive artillery and mortar system." 1/ How will the synergism of combined arms battle affect attrition? Will tanks and aircraft employed independently produce the same effect as when employed in unison?

- Suppression--the temporary reduction of an individual's ability to observe, fire at, or maneuver against an opponent. How does suppression affect attrition? Do all weapons produce the same suppressive effects?

- Logistics--Will ammunition shortages affect attrition? Will fuel shortages have the same effect? And are all types of units, say infantry and armor, equally affected?

But we emphasize that there is more to warfare--and to the models--than "bringing fire upon the enemy." An obvious contrast is the function of command, control, communications, and intelligence (C^3I). Is the strategy to annihilate the enemy, or to seize/defend critical objectives? How might this choice affect the conduct of battle? And, will the modeling of these tactical decision processes, target acquisition, and maneuver be any less important, any less complex than "bringing fire upon the enemy"?

Theater-level warfare is an amalgamation of many combat functions--functions as diverse and yet interdependent as intelligence and casualty treatment. And, while military planning may necessitate the in-depth examination of any one of these functions, it is impractical to put all the detail into a single model. As a result, the models differ--individually, they are strong in some aspects and weak in others. (CEM, for example, details battlefield casualty treatment; the others do not.)

Models are tools. To ask "Which is the best model?" is no more relevant than asking "Which is the best hammer?"-- a watchmaker and a carpenter may have very different opinions.

1/See Dondero [25], p. K-5.

THE VULNERABLE FLANK--THE MANAGEMENT OF QUANTITATIVE TOOLS IN DEFENSE DECISION

> "Do judgment and experience have no place in this approach to the choice of weapon systems and strategy and design of the defense program? Quite the contrary. The suggestion that the issue is judgment versus computers is a red herring. Ultimately all policies are made and all weapon systems are chosen on the basis of judgments. There is no other way and there never will be. The question is whether those judgments have to be made in the fog of inadequate and inaccurate data, unclear and undefined issues, and a welter of conflicting personal opinions, or whether they can be made on the basis of adequate, reliable information, relevant experience, and clearly drawn issues. The point is to render unto computers the things that are computers' and to judgment the things that are judgment's. In the end, there is no question that analysis is but an aid to judgment and that, as in the case of God and Caesar, judgment is supreme." 1/

To repeat an earlier argument: we trust a decisionmaker's judgment; but when that judgement is "extended" by a model--a model that uses unverified assumptions that go beyond science and "objective" fact--"How can the decisionmaker be sure that the model is, in fact, serving as an extension of his/her own judgment?" Where are the checks and balances, the safeguards?

A need for responsible management

Theater-level combat models propose to intervene in Defense Policy. They demand most stringent management control. Yet within DOD, no institutional provision exists for their periodic, comprehensive, and effective appraisal. 2/ If the

---

1/See Alain C. Enthoven, in Tucker [87], pp. 143-144, [emphasis supplied].

2/DOD made commendable, but ad hoc, efforts to address this problem in the early 1970s. See Dondero [25]; Honig [50]; Karr [52]-[56]; Shubik and Brewer [77]; and Walker [107]. Unfortunately, control receives little attention unless one's models do not work--and prudence dictates that more than ad hoc efforts are needed.

findings of this and prior reports are indicative, the present situation is inadequate. 1/ A few brief examples will bring this point home.

There is no documentation for:

- the LULEJIAN model used by the Office of the Secretary of Defense; 2/
- CEM's ground-to-ground suppression module; and
- IDAGAM's refined "kill potential".

Without documentation, there can be no appraisal. And, while it is comforting to imagine that a computer model is mathematically correct and matches the real world, this is not always the case:

- The ATLAS model (documented in 1969) led to the development of the Ground-Air Campaign Model I (GACAM I) in 1971. 3/ The close air support delivered by either side had no effect on the outcome of the battle (as determined by the ATLAS or GACAM I formulation). The point is that both models were documented and running--but until a 1973 appraisal, the close air support was mathematically canceling itself out.

Moreover, incomplete documentation can transform a "visible" aid to judgment into an unquestioned producer of battle outcomes.

- IDAGAM offers the user a choice of several attrition equations at various points in the air combat module--and each equation embodies different underlying assumptions. "Yet the user receives no guidance for making

---

1/Deficiencies in the management, use, and/or documentation of DOD computer models have also been cited in our prior reports: see U.S. General Accounting Office [91] and [93].

2/The OSD version of LULEJIAN is an extensive revision of LULEJIAN I [63]. The analyst informed us that the changes have not been documented except for his personal notes.

3/For model documentation on ATLAS (A Tactical, Logistical, and Air Simulation), see Kerlin [58]; for GACAM I, see Bracken [11].

his choices, no mathematical or physical comparison of the assumptions themselves, and no instructions for making empirical comparisons." 1/

- IDAGAM's results are sensitive to the manner in which a weapon allocates its fire among different target types. Since there is no authoritative reference for this type of information, the data base must use human judgments. Does the decisionmaker care: what battle scenario those judgments pertain to; who made them; and how they relate to the present problem? If the model and the data are not documented, how will he know?

A "black box" cannot extend judgment.

- "After about two turnovers in a military headquarters, or in an analytic organization, you have people [analysts] who were not there when the thing [model] was designed, who don't fully appreciate all the short cuts and approximations that were taken." 2/

- "Where do we find the raw data? Well, the raw data are found in classified and unclassified documents. There are some, but precious few automated sources, and there are a lot of very subjective analysts who are willing * * * to give us their guesstimate of the data. We then take that [data], subject it to human analysis, try to get it into a machine-readable format within the time permitted, and then input it eventually into the model. However, we still wind up with only low level data." 3/

The message is simple and clear. The rationale that says "Defense decisionmakers do not have time to understand and manage policy assisting models" is a direct contradiction to the literal justification which argues that Defense Decision uses such models to compensate for the inadequacies of military judgment. The ethical burden is unequivocal.

1/See Karr [54], p. 45.

2/See Office of Naval Research [69], p. 85.

3/Ibid., p. 211.

A necessary first step

> "Adversary analysis can be useful only to the extent it is open, explicit, and based on common tools and premises. Competing models, each looking at selected parts of the problem, do not provide competing insights, but a competing lack of them." 1/

Richard Steadman's Report to the Secretary of Defense on the National Military Command Structure noted that the studies and analyses which form the basis for recommendations in key areas of policy, strategy, and force planning often have differing results due to wide divergence in models, assumptions, approaches, and computer applications. He emphasized that the studies and analyses conducted to support the Planning-Programming-Budgeting System

> "* * *would be more useful to the Secretary of Defense and the JCS if some proceeded from a common focus, while insuring that dissenting views are expressed. * * *When disagreements arise on assumptions or data, they should be identified and the rationales for the opposing views made explicit." 2/

We agree. The essence of a policy assisting model (or systems analysis) is to provide the framework which permits science and the judgment of experts in numerous subfields to be brought together--made explicit--and used to enhance and extend a decisionmaker's judgment. 3/ If policy assisting models are to be used in Defense Decision, and if the Secretary desires open, explicit analysis so that all parties can meaningfully review and challenge assumptions, then it will be necessary to employ a consistent model (or models) in decisionmaking. This is the theoretical aim of the methodology, and the Nation's best protection against error.

1/See Cordesman [23], p. 194.

2/See Steadman [78], pp. 44-45.

3/A policy assisting model can also be used as an aid in determining research priorities and "guidelines for the development of data collection plans (i.e., what data are important, how accurate they must be, etc.)." See Bonder [9], p. 77; and p. 22 of this report.

This is not to advocate the creation of a "universal" model; nor to imply that individual studies should be confined to a single model. To the contrary, the relevant point is that Steadman's remark pertains not to a modeling problem, but rather to the fact that DOD management does not provide a consistent analytical framework for PPBS decision-making.

Generally speaking, the problem is that while Defense Decision must focus on overall force requirements and inter-Service trade-offs, each Service's POM is developed independently--using different models, assumptions, and data; and drawing on different forms of military expertise. Meeting National Defense needs economically and efficiently depends on achieving a balanced joint capability. Yet, no analytical means of reconciling these differences exists under the present structure.

For each inter-Service trade-off issue, we believe it would be beneficial for DOD management to prescribe a consistent analytical framework of models, assumptions, and data. The framework would facilitate the amalgamation of science and the best judgment from the affected Services. When a party disagrees with some aspect of the baseline, excursions using different data, assumptions, or models could explore the dissenting view--openly and explicitly. To do this, two things are required for the policy assisting models used in Defense Decision:

1. a program of configuration management, 1/ and

2. the creation of a "reference" data base.

There are important initiatives underway which could contribute to this effort. The attempt to establish a VECTOR-2 User's Group is one such effort. When established,

---

1/Applying technical and administrative direction: (1) to establish a documented baseline configuration of model logic; and (2) to control the approval and implementation of subsequent changes to that logic. This is not to preclude local variations of models for separate purposes; but rather to facilitate prescribing a consistant analytical framework on selected studies.

this group will be a cooperative, joint Service effort intended to insure the orderly use of the VECTOR model. 1/ Its aim is to coordinate an exchange of technical information, as well as the planning and costs of future model development.

But data is also a problem. Current data development is ad hoc, time consuming, and expensive. Just in the area of weapons performance data, for example:

> "The amount of data available for each weapon is a function of the availability of the weapon, test specimens, priorities, funds, and development status. Unfortunately, no single source for weapon data exists. The best sources for data are the laboratories, arsenals, and testing facilities which are charged with the development of a particular weapon, munition, or weapon system. For foreign data the best sources are the Defense Intelligence Agency, the Army Foreign Science and Technology Center, the Air Force Foreign Technology Division, and the Army Missile Intelligence Agency." 2/

And here we have a second initiative which promises not only the start of a "reference" data base, but also a considerable savings in the time and cost of both data collection and the preparation of inputs for the various models. That is, the Command and Control Technical Center's (CCTC's) ongoing effort to develop a Weapons Performance Data Base which will--for the first time--make input data for the models available to common users from a centralized location.

A true "reference" data base, however, would have to include all of the input data required by the models--not only weapons performance data, but also data pertaining to tactics, terrain, supply, and order of battle. This information is required for the National Military Command Center's models and for the models used at numerous Service activities throughout the United States. It makes sense that substantial benefits could be derived through the establishment

---

1/Participants in this effort include the Studies, Analysis, and Gaming Agency; the Defense Communications Agency's Command and Control Technical Center; the U.S. Air Force Studies and Analysis; the U.S. Army Concepts Analysis Agency; and the U.S. Army Harry Diamond Laboratories.

2/See Command and Control Technical Center [20], p. 3. For further discussion: see Office of Naval Research [69], pp. 208-244.

of a central focal point for the collection, management, and quality control of these common data requirements.

Both initiatives--the VECTOR-2 User's Group and CCTC's Weapons Performance Data Base--appear to offer economical improvements in the management of quantitative tools in Defense Decision. But these are small efforts in relation to what is needed.1/ They deserve careful consideration for extended applications.

1/The reader should realize that there is a second, equally important dimension to the concept of a consistent analytical framework. That is, there should be a consistent representation of a weapon system's "effectiveness" between low-level and high-level models--between procurement and force planning decisions.

In other words, improvements in weapon systems analyzed at the low level "must be reflected in the capability of units using this weapon system in a more aggregate sense. That is, a battalion with a much improved tank must be better than battalions with outdated equipment. This improvement must reflect itself at the division level where mixes of better battalions and poorer battalions must be considered. In addition, new equipment and new concepts of operating with different types of battalions must improve a brigade or division capability by a sufficient amount to be recognized in theater-level analyses." (See Honig [50], p. I-9.)

The hierarchy of models approach is one method for achieving this consistency. Here, a family of separate models is employed to represent discrete levels of combat ranging from small unit encounters up to a theater-level conflict. At each level, output from the more detailed lower model becomes input to the next higher model.

The Army has employed this hierarchical concept on a limited scale since the early 1970s; but only to the extent of linking a computer simulation of battalion combat to a division-level war game.

This year, the Army's Special Study Group recommended adopting the hierarchy of models approach as a framework for Army analyses. That recommendation envisioned a hierarchy of models extending from the individual weapon up to theater-level conflict. The Study Group also recommended the adoption of an Army-wide data base system to provide centralized management, access, and review of data requirements. See: Hardison, D.C., study director, Review of Army

RECOMMENDATIONS TO THE SECRETARY OF DEFENSE

We recommend that the Secretary of Defense:

- Reassess the adequacy of current practices in the management and use of policy assisting models employed in Defense Decision. This should include

  --identifying corrective measures needed to insure that models in use are well documented, comprehensively appraised, and updated on a formal basis; and

  --insuring, on a continuing basis, that the models are employed in a manner consistent with their theoretical aim, which is to enhance and extend the decisionmaker's judgment.

- Develop procedures to enhance the policy assisting models' contribution to open explicit analysis in the key areas of policy, strategy, and force planning. Such procedures should recognize the economies and efficiencies offered by

  --implementing configuration management for the models used in Defense Decision;

  --establishing a "reference" data base to support multiple users; and

  --prescribing the model(s) and data to be used in the analysis of specified issues.

AGENCY COMMENTS AND OUR EVALUATION

The Department of Defense was given the opportunity to review and comment on this report. In its letter of September 24, 1979 (see app. IV), the Department told us:

Analysis, Volume I-Main Report, Special Study Group, Department of the Army, Washington, D.C., Apr. 1979.

These recommendations portend long-term efforts. They have been accepted in principle, and are undergoing further study and review. We mention them here to demonstrate an awareness of the need for a consistent analytical framework at the Service level.

The apparent "belief that decision makers at the highest level of the Defense Department depend heavily on the results of specific models when making decisions is erroneous. Models are employed within the framework of studies, with the intimate workings of the model seldom going beyond the study director. The results of the model are interpreted and evaluated with relationship to the study. The degree to which the results of the model are incorporated in the study report is usually dependent upon the study director's judgment of the validity of the model and the confidence that he has in the results. Models in themselves seldom enter into the higher echelons of decision making; studies do."

The Department also pointed out that:

- "Presently the Department of Defense does specify what models and what data are to be used in specific studies. However, this technique is not, and should not, be used in every study. In many cases there is no standard, prescribed method of solution. In such cases the DoD component conducting the study should have the responsibility and freedom to develop the necessary methodology, models, and data to satisfy its requirement."

- "Measures toward '...establishing a reference data base to support multiple users' are already being implemented. The Office of the Assistant Secretary of Defense (Program Analysis and Evaluation) is presently updating the Defense Force Planning Data Base. This will assist in solving many of the problems associated with data collection efforts identified in the report. It will provide users with a consistent source for input data."

- "Department of Defense Directive 5010.22 specifies the policy to be followed with respect to studies and analysis. Specifically, it requires that a report associated with each study document all information on assumptions made, models used, model modifications, effectiveness measures, criteria and sources of intelligence or data used."

The Department did not address the recommendation about implementing configuration management for the models used in Defense Decision.

We fully support the Department in noting that studies, not models, enter the highest echelons of Defense decision-making. As we said earlier:

> "No single model or analysis drives an individual decision. Rather, groups of studies tend cumulatively to foster a consensus on major issues. In doing so, the analytical efforts interact strongly with external realities and decisionmaker's opinions" (p. 46).

We also support the Department's other comments--in their specifics. But to generalize from those specifics is to obscure the very weaknesses that our recommendations seek to correct.

Let us emphasize at the outset that this report intends no appraisal of specific models and studies, nor any criticism of specific decisions. We are, in a sense, looking at the "methodological efficiency" of the analytic structure supporting Defense Decision. What we have attempted to illustrate thus far is that the Services, JCS, and OSD(PA&E) employ different models, different data, and different forms of expertise in developing force structure and force level recommendations for the Secretary of Defense. 1/ That this arrangement provides a variety of perspectives is without question; but we should be under no illusion that variety is a full and sufficient criterion for quality decision support.

Quantitative methodology has considerable potential to aid in the analysis of squishy problems; but that potential rests in its ability to serve as a framework for bringing together the "best" of science and considered judgment. If relevant judgments are omitted from the framework, the benefits of the methodology are diminished (suboptimized). Let us note, therefore, that a model--such as CEM--is a framework. Two models--such as CEM and IDAGAM, or two versions of CEM, or even the same version of CEM with two different sets of data--constitute two different frameworks. In short, analysis is unique to its framework. To go between frameworks is to bridge an analytic discontinuity involving far more than just

---

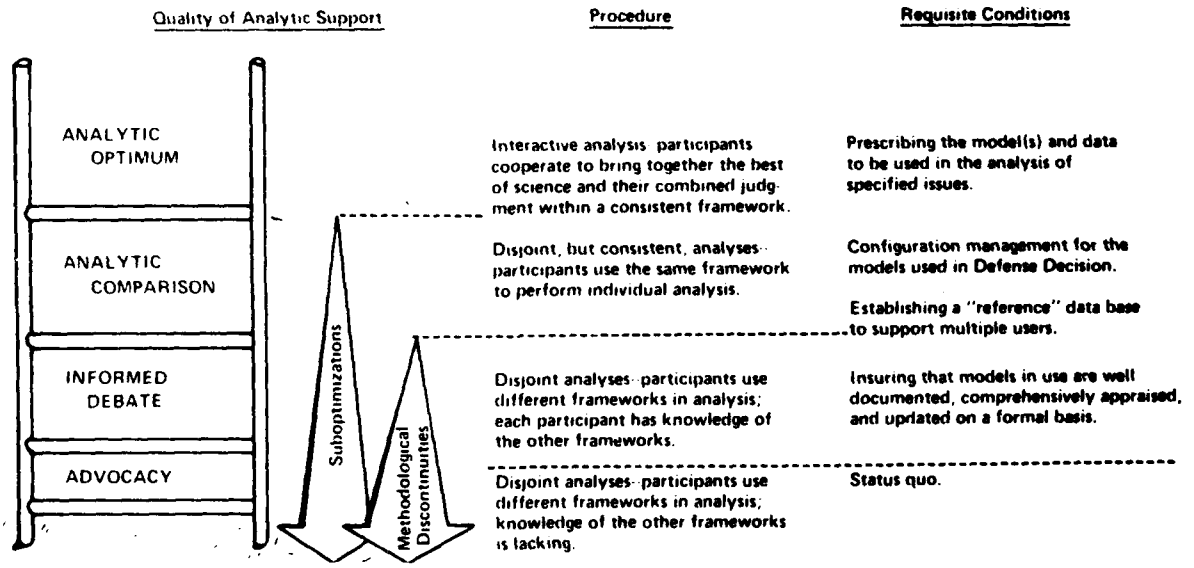
1/Our discussion has naturally entailed gross simplifications about the range and scope of models, issues, and organizational entities involved in Defense decisionmaking. The inclusion of real world complexities would not change our basic message.

mathematics (e.g., tank vs. aircraft attrition may be compared in CEM's combat, but it would be a pitfall to believe the methodology supports meaningful comparisons between CEM's tank and IDAGAM's aircraft attrition).

If we look closely at the effects of these suboptimizations and methodological discontinuities, we can briefly identify four rungs in the "decision support" ladder. From the top down, we term these: the analytic optimum, the analytic comparison, informed debate, and advocacy. Figure 4.13 merely outlines the basic concepts and principles of this "ladder."

Figure 4.13

The Decision Support Ladder



As our figure attempts to illustrate, the quality of analytic support provided for high-level Defense decisionmaking depends on more than the expertise of individual analysts; it also depends on the "methodological efficiency" of the Department's analytic structure. In other words, for analysis in support of decisionmaking:

- A single framework and integrated judgments can provide the optimum analytic recommendation.
- A single framework and independent judgments can permit the analytic comparison of differences in judgment between recommendations (e.g., the framework is consistent, only the judgments change).

- Different frameworks and independent judgments confound between-recommendation comparisons (e.g., all things change, differences reflect variations in framework and judgment).

This is not an argument against examining various aspects of a decision problem; it is simply a statement that the sacrifice of a consistent analytical framework on central issues may counterbalance any moderate advantage in variety.

The Department has stated that it does specify what models and what data are to be used in specific studies. We agree; and we point out that, in general, such studies represent the "analytic optimum" rung on the figure's "decision support" ladder. But we also realize that practical constraints make it necessary for the Department to employ this procedure in a highly selective manner. The majority of analyses supporting PPBS force planning and inter-Service trade-off decisions fall below this rung.

The Department's second point, that "Measures toward '. . .establishing a reference data base. . .' are already being implemented," is a positive step toward strengthening the second rung of our "decision support" ladder. We are told that this effort focuses on static indicators (see p. 148) and on improving the consistency of analyses performed internal to OSD. 1/ We believe this consistency is, in itself, important. Nonetheless, consistency within OSD analyses is one thing; and a consistent framework between OSD and the Services' analyses, quite another. The majority of analyses supporting the PPBS force planning and inter-Service trade-off decisions fall below this second rung.

1/The Defense Force Planning Data Base originated with the NATO Task Force Action Memorandum of 16 August 1973. Until recently, its development has proceeded on a catch-as-catch-can basis. As a result, the present data base has problems--involving the acquisition and quality control of data, the lack of a common counting or estimating methodology, the inability to simultaneously update data, etc. The Department has begun a major effort to alleviate this condition. Its sustained efforts will, in our opinion, enhance the information available to OSD and the JCS for balance assessments and posture analyses.

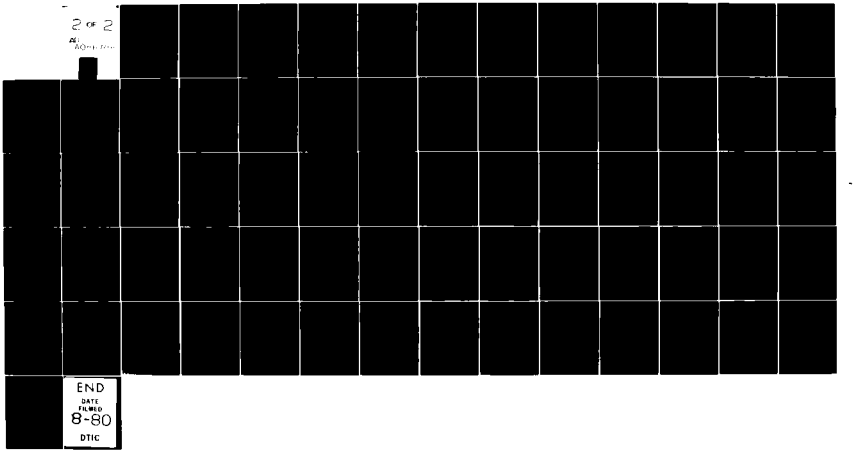
The third rung is entitled "Informed Debate." Here the Department points out that DOD Directive 5010.22 requires each study to "document all information on assumptions made* * *." That the Directive is an excellent policy statement is without question, but we are less sanguine about its enforcement. A trivial example of high-echelon decision support will serve to illustrate:

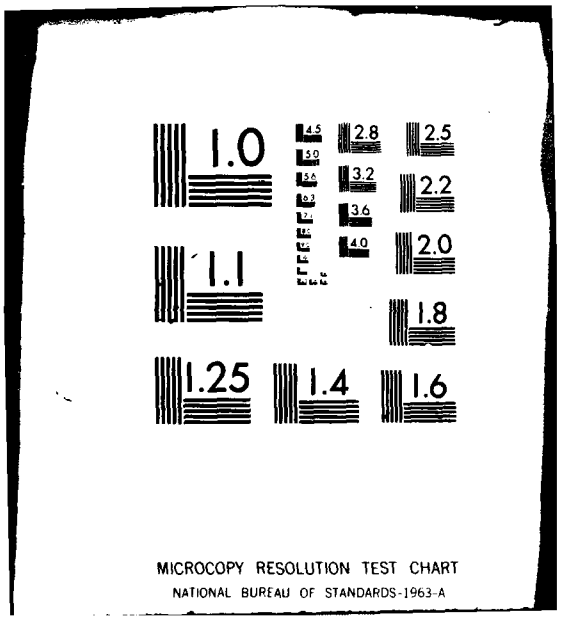
> "For determining how many tanks we should have we conducted a war reserve study on tanks. For that study, P&E used a Lulejian model. The analyst completely controlled it, ran it from beginning to end, and used it in a sensitivity analysis to decide how many tanks were needed in war reserves." 1/
>
> This study was presented as an Issue Paper for high-echelon Defense Decision. Three points are worth noting: (1) the model, as we pointed out on p. 74, was not documented; (2) the analysis, solely dependent on the judgment of PA&E's analyst, was a suboptimization; and (3) the difference between this model and those used by JCS and the Service (see pp. 64-66) introduced a methodological discontinuity between information supporting requirements determination and budgeting. The example is by no means unique. 2/

1/See Office of Naval Research [69], p.17.

2/For additional examples of problems of incomplete documentation, suboptimization, and methodological discontinuities, the reader is referred to: Office of Naval Research [69], pp. 14, 26-27 (examples concerning PRM-10); Hardison, D.C., study director, Review of Army Analysis, Volume I-Main Report, Special Study Group, Department of the Army, Washington, D.C., Apr. 1979, p. 10-1 (concerns about "a lack of information flow among models run by different agencies"); and Shupack, S. L., An Examination of the Conceptual Basis of the Attrition Processes in the Instititute for Defense Analyses Ground-Air Model (IDAGAM), Naval Postgraduate School, Monterey, California, Mar. 1979, pp. 10 and 14 (discussion of the IDAGAM model's use). Further examples may be found in many of this report's other references--including Cordesman [23]; Honig [50]; and Steadman [78].

AD-A086 708 GENERAL ACCOUNTING OFFICE WASHINGTON DC PROGRAM ANAL--ETC F/G 15/3
MODELS, DATA, AND WAR: A CRITIQUE OF THE FOUNDATION FOR DEFENSE--ETC(U)
MAR 80
UNCLASSIFIED GAO/PAD-80-21 NL




1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
6.3
7.1
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

We believe that DOD has nurtured a competent staff of analysts who are, in general, doing high quality work. Nevertheless, the Defense decisionmaking process is sufficiently complex that it is extremely difficult to support high-echelon decisions with single analyses, each embodying the "best" of science and relevant judgments. We would expect, therefore, that the Department would be strongly inclined toward alleviating the effects of suboptimizations in, and methodological discontinuities between, these analyses. The critical problems, in our view, are not technological, they are institutional. Improving the "methodological efficiency" of Defense Decision is the theme of this report--and the purpose of this chapter's recommendations.

CHAPTER 5

EMPIRICAL STUDY IS NECESSARY TO STRENGTHEN THE SCIENTIFIC FOUNDATION AND OBJECTIVITY OF DEFENSE DECISION

> "In the context of a structured studies and analysis program, the application of quantitative methodology can be of great help. However, it must have a richer and healthier empirical foundation than it now has." 1/

The models used to support Defense Decision need to be more transparent to the decisionmaker, but achieving this goal will be a long-term effort that must be deliberately initiated and supported by management.

Currently, basic combat phenomena--attrition, breakpoints, suppression, synergism, FEBA movement, 2/ etc.--are not well-understood. And much of the data used in the analysis of conventional force requirements rests on a shaky empirical-theoretical foundation.

> "The relevance of increments of weapon technical performance to actual combat utility is at best obscure, if not unknown, because of inadequate testing to uncover whatever relationships may exist between measures of combat merit and measures of technical performance." 3/

In relation to a NATO scenario, for example, we cannot quantify objectively:

- The true battlefield impact of the XM-1 versus the M60A3 tank;
- How the total tank requirement will be affected by the introduction of precision guided munitions; or

---

1/See Stockfisch [80], p. 4.

2/FEBA movement is a primary metric of battlefield success in theater-level models. In the aggregated models "the rate of advance of the FEBA is assumed to be a function of the force ratio...It seems quite likely that this is far too simple. The rate of advance is surely affected just as strongly by other factors, especially the mission." See Honig [50], p. VI-12.

3/See Stockfisch [80], pp. 10-11.

- The cost-effectiveness trade-offs between tanks, armed helicopters, and attack aircraft.

Since there is no generally accepted or independently validated "theory" of combat, this condition probably cannot be entirely avoided. But it is not an acceptable or satisfactory condition, and it should be of legitimate concern to all those involved in Defense Decision.

A necessary first step to put the methodological house in order is for Defense decisionmakers to understand that studies and analyses used as management tools require adequate maintenance and support. However, simply to make more resources available for additional empirical work is not sufficient. Considerably more attention, for example, is needed to detail carefully the limitations of existing theory and data. This list is large and well exceeds the scope of this report. In fact, a reasonable task might be to determine what the individual components of that list are in terms of building research and policy priorities aimed at developing knowledge, understanding, and verified formal structures (theories) about the combat phenomena that are integral parts of warfare. 1/

Although many combat models represent these processes, their mathematical logic is without substance. "We do not have sweeping laws such as F=ma and E=IR. In many instances, we cannot even define the F's, m's, and a's." 2/ Moreover, experimentation designed to develop such fundamental knowledge is scarce. What then is cost-effectiveness? Certainly not objective science. In our view, a reasonable program of historical research and field experimentation--adequately directed and controlled--is essential for the long-term integrity of Defense Decision.

1/"It is clearly worth noting that mathematics research on optimization, stochastic processes, statistics, and others, is significantly different in substance and intent from that being proposed. Mathematics research omits of any necessary connection to real world phenomena. It produces 'tools of the trade' with no integral content, just as differential equations have no integral content regarding the motion of a pendulum. In contrast, the focus of the research proposed is substantive, with the intent of describing real world operational and management phenomena and developing associated causal dynamics." See Bonder [10], pp. 18-20.

2/Ibid., p. 20.

HISTORICAL RESEARCH: A PLACE TO START

Historical research is the key to the lessons of history. 1/ While reconstructing battles from incomplete records is difficult, and while some may question the relevance of history to the modern battlefield, let us point out that science is based on repeatable observation. We doubt that historical analysis will provide definitive, quantitative values for today's battlefield; but, if properly done, such an analysis should reveal persistent patterns--indicators of which combat processes are most important and warrant detailed investigation by other means. If there are no patterns, then there is no "quantitative theory of war," in which case we cannot assist judgment with quantitative tools.

> "That the historical record is the place to go for a base from which to project, it is necessary only to point out that this is precisely what has always been done. * * * The trouble is that the basis established in the past has been inadequate." 2/

What judgments rest on the "pace" of battle?

The rates of advance which control FEBA movement in the aggregated models "were originally developed on the basis of limited historical data. The rates have been subsequently modified, changed, and aggregated so that current rates of advance have little, if any, traceable connection with historical fact." 3/ (Appendix II traces this transformation in greater detail.) Subsequent research efforts on FEBA movement/rates of advance--studies without any overall coordination or direction--have produced inconclusive, often contradictory, findings. 4/

1/Not all of history's lessons pertain to weapons. The Historical Evaluation and Research Organization, for example, has found interesting relationships between mobility and dispersion, and between weapon lethality and deployment. See Dupuy [28], pp. H-28 - H-33.

2/See Dondero [25], p. G-20, [emphasis supplied].

3/See Honig [50], p. II-15.

4/See Dondero [24] and [25]; Dupuy [29]; Goad [36]; Graves [37]; and Zimmerman [115]. A point worth noting is that historical research is time consuming. As a result, available data bases are limited in scope and magnitude--and data limitations can produce inconclusive findings in any analytic effort.

Simply stated, existing movement rates are based on tenuous data. "In fact, there are major differences--involving factors of three to five--in the values used by the different NATO nations, and even by organizations within the same nation." 1/ These differences produce widely varying estimates of the length of time during which conventional warfare operations can be maintained--an important consideration in studies involving critical, real-time related issues. In a NATO scenario, for example:

- How long will NATO forces have to prepare defensive positions?
- What is the "pace" of battle--how many soldiers will be needed, and how fast? How long will the Nation have to mobilize, and what demands will this place on the Reserve, National Guard, and Selective Service System?
- What NATO airfields will be available for the airlift of strategic reserve forces?

These concerns are nontrivial--the answers to such questions have a profound impact on the formulation of national policy. Improvements are needed in the management and coordination of historical research.

What judgments rest on attrition theories?

Engel's classic verification of Lanchester's theory using Iwo Jima combat data produced a remarkable fit (see figure 5.1). 2/

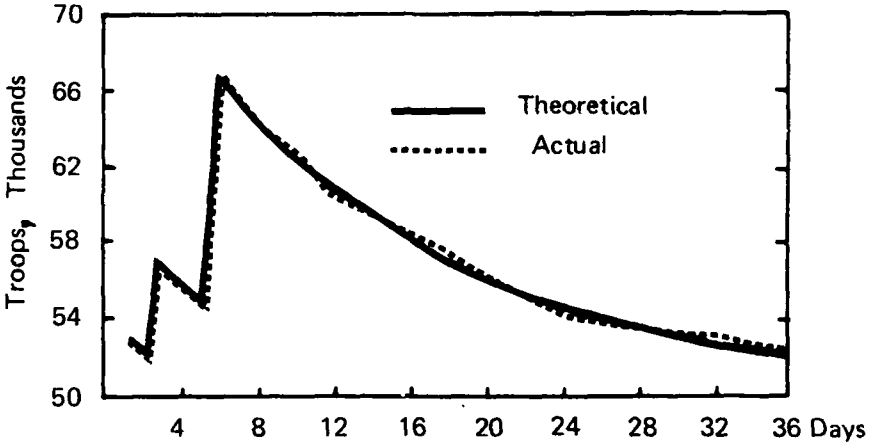
1/See Goad [36], p. 1.

2/See Engel [30]. Let us also note that Dr. Engel has cautioned against the conclusion that his work validates Lanchester's theory; see Engel, in Fain [32], pp. 5-7.

Figure 5.1

Theoretical and Actual Survivors During the Capture of Iwo Jima

(Engel's Verification of Lanchester's Theory)



Source: Operations Research 2,
"A Verification of Lanchester's Law," 1954.

But while this study is a classic example of historical verification--the fact is that its results are widely contradicted by numerous other studies. 1/ Lanchester's theory appears very sound for homogeneous forces engaged in combat on the "playing fields of Eton"; but when heterogeneous forces, undulating terrain, battlefield obscuration, etc., must be considered--Lanchester's simple theory no longer applies. There is no validated theory that provides the shape of the attrition curve in these scenarios.

But the Nation must plan and budget. And to do so, decisionmakers must render judgments that are inextricably tied to attrition considerations--judgments that affect the Federal budget in diverse ways. Consider, for example, how estimates of ammunition expenditures and personnel/equipment casualties in a NATO scenario might affect DOD's requirements for:

1/For other verification attempts: see Busse [14]; Fain [31]; Helmbold [41]-[46]; Ostermann [70]; Overholt [72]; Samz [76]; Weiss [109] and [110]; and Willard [112].

- War reserve ammunition,
- Military hospital construction, and
- Airlift/sealift resupply.

We trust a decisionmaker's judgment, but a distinction must be made between intuitive judgment and scientific understanding. On such matters of national policy, it seems prudent to underpin decisionmaking with the best possible empirical support.

FIELD EXPERIMENTATION: A PLACE TO FOLLOW UP

Field experimentation is the essential link between history's insights and a "science" for the modern battlefield. There are several types of field experiments, but for the sake of simplicity, let us say that we are talking about highly realistic mock battles conducted on instrumented ranges--ranges that use electronic equipment, lasers, etc., to collect and record real-time data on selected combat phenomena (e.g., movement rates, firer-target locations, rounds fired, hits and near misses). The objective is to provide a "realistic," empirical assessment of how weapons, tactics, and other combat phenomena affect battle.

The U.S. Army has over two decades of experience with field experimentation. That experimental capability has grown in response to the need for better evaluations of new weapons and equipment. But it has also enhanced--and sometimes corrected--our understanding of elementary combat phenomena. For example:

- The Suppression Experiment (SUPEX) analyzed the proximity of fire required to suppress a given threat, the volume of fire required to maintain a specified level of suppression, and the suppressive effects of selected weapon systems. 1/ The experiment provided important insights--but there is much more to be learned.
- A comparison of the TETAM (Tactical Effectiveness Testing of Antitank Missiles) Field Experiment with predictions of CARMONETTE (a high-resolution model)

1/See U.S. Army Combat Developments Experimentation Command [88].

> indicated that the "digitized LOS [line-of-sight] models used in our Monte Carlo simulations of battalion-level combat produced significantly erroneous LOS realizations." 1/

That the Army's experimentation program has provided much useful information is without question. It should also be understood that the existing program is neither instrumented nor intended to examine the phenomenology of battalion-level combat.

A data base for the phenomenology of combat

After having said this, let us point out that the President's FY 1980 Budget provides initial funding for a National Training Center (NTC). That Center, supported by the Secretary of Defense, the Army, and the Defense Advanced Research Projects Agency (DARPA), is intended to provide a fully instrumented, realistic combat environment for training U.S. maneuver battalions in combined arms operations against a suitably trained and equipped aggressor force. Now, this is a key point. The maneuver battalion is the basic tactical unit of land warfare. It is the force that coordinates and applies the combined firepower of infantry, armor, artillery, and air. And the NTC, instrumented to support the training and debriefing of these units, will collect and record a detailed chronology of each engagement's:

1/The quote is Bonder's [9], p. 84. For a description of the experiment: see Thorp [86]. There are indications that LOS realization is only one of several major problems in the present methodologies for interfacing terrain, tactics, and weapon systems in COEA and force design studies. "There is extreme sensitivity in combat model results as the scenarios (terrain and movement assumptions) are varied, even when this variation is within a class of scenarios chosen for their a priori equivalence." This sensitivity is such that "weapon system or force design choices may be reversed between equivalent or equally likely terrain and scenario choices." See Farrell and Freedman [33], p. 38, [emphasis supplied].

- Battle Management
  - --Decision processes, orders, reasons
  - --Intelligence reports, sources
  - --Communication delays
  - --Fire orders, response times, accuracies
- Tactical doctrine
  - --Sub-unit location, formation, movement rate
  - --Target detection, actions
- Weapons employment
  - --Firer-target locations, intervisibility
  - --Hit/kill probabilities

What we are saying here is nothing less revolutionary than that the Army has provided the "know-how" and technology --and the NTC will potentially provide the "maneuver-battalion data base"--to begin a scientific inquiry into the phenomenology of combat.

The message is simple and clear, if one accepts these premises:

1. That tactics can be taught--implying that the outcome of battle is not an arbitrary, random process;

2. That the maneuver battalion is the basic tactical unit of land warfare;

3. That the things to be learned in the NTC's "surrogate combats" have applicability to war; and

4. That quantitative tools aid decisions on weapon systems, force composition, and force levels.

Then, logically, one must also accept the premise that the NTC's data base will be an important resource for Defense Decision. And, while we should be under no illusion that the NTC's data base will provide a precise rendition of battle, it should be able to answer--with empirical support and confidence levels--some very practical, "policy relevant" questions. Questions that deal with such fundamental phenomena as:

- The attrition curve. What is its shape? How is it affected by variations in tactics, terrain, weather, and visibility (day/night actions or battlefield obscuration)?

- The structural distribution of attrition. Who and what are most likely to be attrited--tank gunners or infantrymen, artillery or mortars? What skills and equipment are most critical to the battalion's performance? 1/

- The value of Close Air Support (CAS). How does the level, type, and timing of CAS affect battalion-level battle? Which has the greatest impact on unit performance--the aircraft's ability to inflict attrition or its ability to disrupt and disperse the enemy (thereby enabling the battalion to accomplish its mission quicker, and with fewer casualties)? What is the optimum balance between attack helicopters and fixed-wing aircraft? between response-time, loiter-time, and ordnance payload? 2/

The idea of providing a coherent framework for training and analysis is strikingly simple. It ameliorates one of the most debilitating challenges to Defense Decision--the perishability of knowledge. Army training and Defense planning are both oriented toward a future conflict that bears little resemblance to any fought in the past. Even the conditions under which that combat will be fought are constantly changing as advanced technologies and tactical innovations combine and build upon each other. The mass bombing raid--an innovation of World War II--is an anachronism today. The mortar, a lethal weapon against foot soldiers in past conflicts, may have a very different role on the mechanized battlefield.

1/See Clark [16], p. 34. "The statement that a unit can be considered no longer combat effective when it has suffered a specific casualty percentage is a gross oversimplification not supported by combat data." Also see: Aldrich and Bode [1]; Best [5], [6]; and Stockfisch [79].

2/CAS has been a subject of continuing interest to DOD for many years. The literature abounds with mathematical studies that evaluate "CAS effectiveness" in terms of an aircraft's ability to destroy enemy targets. Alternatively, had those studies evaluated "CAS effectiveness" in terms of the aircraft's passive role (presenting a visible threat; modifying enemy behavior; increasing the effectiveness of ground force weapons/tactics), their implications for aircraft design and CAS doctrine might have been very different. We found few works of empirical research on this topic. We were unable to identify an appropriate measure of "CAS effectiveness." The NTC may provide some interesting insights on this issue.

So, "How does one obtain the empirical relationships needed for military planning?" We know of no better way than by using field experiments and the NTC as a link between the insights of history, and a "science" for the modern battlefield.

EVOLVING A "THEORY" OF WAR: A PLACE FOR TOP MANAGEMENT INVOLVEMENT

The accuracy with which Defense requirements are determined exerts a powerful influence on our Nation's security, and economy. Yet only as war is understood can the reality of these Defense requirements be judged, and the necessary resources matched with needs.

This operational imperative--understanding war, not only as abstract "principles", but also as empirical propositions with precise quantitative content--has been emphasized by Alain Enthoven. His argument is persuasive:

> One need only examine the "history of our deliberations over NATO strategy to discover that our inability to assess the capabilities of General Purpose Forces (on both sides) can, in fact, have potentially dangerous consequences."
>
> In the 1950s, "the belief was widespread in the West that the Soviet Union had something on the order of 160-175 divisions in its Army, * * * each of which, it was assumed, had an effectiveness roughly equal to a NATO division. * * * This strength ratio of nearly 10-1 [in favor of the Soviets] made the job of effective non-nuclear forward defense seem hopeless. It is not surprising, therefore, that the United States Government in the 1950's fell back on the strategy of 'massive retaliation'--that is, the early use of strategic nuclear weapons, even in the case of limited non-nuclear attack--because there seemed to be no possibility of an effective alternative. Whatever else one might say about this strategy, it certainly entailed major risks. * * * In fact, it has turned out that the true picture is very different from this pessimistic one." 1/

Much progress has been made since Enthoven's remark. Today, the determination of requirements and the allocation

---

1/See Enthoven, in Tucker [87], p. 187, [emphasis supplied].

of military resources--both within and between the Services--is centrally managed by the highest levels of Defense Decision. This arrangement is supported by the Planning-Programming-Budgeting System (with its analytical engine, cost-effectiveness analysis), and it is directed toward realizing one objective: to get the most defense out of any given level of resources or, what is logically equivalent, to achieve a given level of defense at the least cost.

But while the economic theory can be simply stated, its reduction to practice is a much more difficult matter.

- How many divisions of what type are required to hold the line in Europe? Does this depend on Airlift/Sealift reinforcement rates? If so, how might that number be affected by the Airlift Enhancement Program?

- Can new artillery technologies (like the COPPERHEAD projectile, with its laser-guided antitank capability) be used as an economical substitute for tanks or close-support aircraft?

- If the procurement justification says the XM-1 tank is more than twice as effective as the M60 tank, does this mean that an Army Division can achieve the same fighting capability with half as many tanks?

- Is the increased accuracy of the DIVADS antiaircraft gun worth the additional cost?

These are very difficult, complex, challenging, and vital questions to which we now have quite imperfect answers. Just in the case of the DIVADS antiaircraft gun, for example, "How important is 'accuracy' to its mission of protecting front-line combat forces?" An interesting perspective on this issue can be drawn from the British experience in World War II:

> "At the beginning of the war, a great number of British merchant vessels were seriously damaged by aircraft attacks in the Mediterranean. The obvious answer was to equip the vessels with antiaircraft guns and crews, and this was done for some ships. The program was a somewhat expensive one, however, since antiaircraft guns were needed in many other places also. Moreover, experience soon showed that single guns and crews, with the little training which could be spared for merchant vessels had very little chance of shooting down an attacking plane.

> Questions concerning the soundness of this allocation of scarce resources were raised when reports showed that the gun crews were shooting down only about four percent of all attacking aircraft. "This was indeed a poor showing, and seemed to indicate that the guns were not worth the price of installation. On second thought, however, it became apparent that the percentage of enemy planes shot down was not the correct measure of effectiveness of the gun. The gun was put on to protect the ship, and the proper measure should be whether the ship was less damaged if it had a gun and used it, than if it had no gun or did not use it. The important question was whether the antiaircraft fire affects the accuracy of the plane's attack enough to reduce the chance of the ship's being hit." The guns were doing the job rather well; of the ships attacked, 25 percent of those without protection had been sunk, while only 10 percent of the ships with protection were lost in the same period." 1/

These differences are really the heart of the problem, and it is not an easy one to solve. What is the effectiveness of a weapon system? How can it be measured? How can it be linked to the broader objective of determining the most effective mix of ground and tactical air forces that can be bought and maintained for, say $30 billion per year? How can these measures of effectiveness (MOEs) be structured to support congruent economic choices across the spectrum

---

1/See Morse and Kimball [66], pp. 52-53, [emphasis supplied].

of DOD resource allocation decisions? How can the MOEs be made compatible between: 1/

1. Selecting performance characteristics of new weapons;
   --What constrains the effectiveness of a weapon's design: military need, technology, or cost?
   --How is the effectiveness of a tank measured and compared to that of a TOW antitank weapon?

2. Allocating resources among specified combat functions;
   --How is the effectiveness of a tank battalion influenced by the enhanced effectiveness of a new tank? Does the battalion require as many tanks?
   --What combination of tanks and TOWs provide the most effective antitank capability for a given cost?

3. Allocating resources among military services; and
   --Is the effectiveness of a tank battalion constrained by Airlift/Sealift capacity; and if so, should we trade-off tanks for additional Airlift/Sealift resources? Or, should we have bought more TOWs?
   --How is the effectiveness of a tank battalion compared to that of an attack aircraft squadron?

4. Determining Theater force level requirements?
   --What MOEs are used in determining "What forces are needed to hold the line in Europe?", and "How do they relate to the performance characteristics of new weapons?"

1/While the integration of cost and effectiveness information is central to the notion of rational economic choice, the theory implicitly assumes that the analysis is fed with adequate empirical data. In this instance, the decision-making hierarchy can be reasonably assured that the MOEs used in lower-level decisions are consistent with those used at higher-levels. This is essential to avoid diseconomies. The diligent reader will no doubt wish to consider the important and subtle implications of not having a rigorous empirical-theoretical basis for prescribing these MOEs --and the challenge it presents to centralized decision-making. We elaborate only that, either explicitly or implicitly, the choice of MOEs bears directly on: the incentives for weapon design; the quantitative-qualitative debate; and the economics of Defense Decision.

Defense Decision--whether based on military judgments, or sophisticated economic techniques--is critically dependent on the knowledge of what a military force can reasonably be expected to do. That measure is essential, not only to permit realistic force comparisons, but also for gauging the incremental contributions of new weapons and tactical concepts.

However, from a scientific point of view, the present "understanding of war"--at least insofar as the effectiveness of conventional ground and tactical air forces is concerned--is in a relatively primitive state. 1/ Basic research aimed at understanding the fundamentals of combat is needed, but quantitative techniques have not been systematically applied to achieve these discoveries. Yet, such discoveries are the heart of the Planning-Programming-Budgeting System. There follows the requisite that top management must be deeply involved. 2/

Much research has been supported by the Services over the years, and the individual work has generally been of high quality. Nevertheless, the overall research program has been without strategic guidance. Top management has not provided research priorities for the development of knowledge, understanding, and verified formal structures (theories)

1/This is not to imply that present judgments are inadequate; but rather that the scarcity of attention given research "could lead one to believe that the DOD is modeling problems that are easily quantified and are well enough understood that no new theoretical research is needed to explain them. If so, then either the problems confronting the DOD are in fact being managed in a rigorous, scientific fashion, or, as appears to be the case, intangible, intractable, or 'soft' issues are largely being overlooked." See Shubik and Brewer [77], p. 17.

2/What is needed is not "a simple answer, a single calculation or index, certainly not a substitute for experienced judgment. * * * We need methods that would enable us, when experts disagree, to develop a rational and explicit process for defining the disagreements and for finding a systematic way to resolve them." See Enthoven, in Tucker [87], p. 195.

about elementary combat phenomena. 1/ As a result, past efforts were structured to satisfy immediate needs--not long-term goals--and, hence, are criticized for being inadequately documented, narrow in scope, and for failing to preserve collected data in detail. 2/ Subsequent studies have often been forced to "reinvent the wheel." 3/

> "It seems reasonably clear at any rate that in 15 years there hasn't been too much significant improvement or alteration of the original constructs of Hulse and Parsons [FEBA movement/rates of advance] * * * But I should like to think that increasing and continuous attention will be given to the systematic accumulation of data from history and field experimentation. Failing reasonable efforts along these lines, we could probably

---

1/This is not to imply that top management does not initiate specific projects and approve annual research programs, but rather that it has never detailed the limitations of existing theory and data, nor provided a list of priorities for coordinating long-range research into the phenomenology of combat.

2/"Additional data which might have been easily recorded or other factors which might have easily been included were not considered because of the shortsightedness of the sponsor and/or experimenter." See Honig [50], p. V-19. There are, of course, exceptions. Experiments conducted by the U.S. Army Combat Developments Experimentation Command appear to be very well documented, and special care is taken to preserve and make available experimental data (see U.S. Army Combat Developments Experimentation Command [89]). The SHAPE Technical Centre's analysis by Goad provides an example from outside DOD. Goad includes his data as an appendix to the analysis (see [36]). Unfortunately, such practices are not universal.

3/"The plea here is for recognition that good data do not die. They generally represent a substantial investment of money and time, and the failure to preserve them for future use is a crime." See Honig [50], p. V-20. But data management, in itself, requires a substantial investment. The ability to discriminate--to determine what should be recorded and what should be preserved--is crucial. Hence, our concern for a coherent overall research strategy.

convene a seminar like this 5 years from now and say the same things." 1/

Eight years have passed since that remark--and in many respects, one might still "say the same things." The need for a basic research strategy spans the entire spectrum of combat phenomena.

A related but distinct point is that this absence of top management guidance also appears to affect the development of required analytic capabilities. A synopsis of representative comments obtained during our interviews with officials from DOD; the Services; Federal Contract Research Centers; and private and consulting organizations, such as the General Research Corporation and the RAND Corporation, will serve to bring this point home.

- Present information is conflicting--we just don't understand the combat processes well enough.
- People often discuss the problems with models, but no one conducts a dialog on the nature of a research strategy aimed at solving some of the problems.
- The search for a theory of combat is legitimate--at least some theory to explain the asymmetries of combat; i.e., to explain in a given situation the effects of one (or several) weapons against another.
- Current research funding is too short-sighted. No group or person is promoting, funding, or even considering long-range research--the type needed to correct many of the modeling problems we have currently. Nearly all research being funded must be conducted in the short-term; i.e., during the term of the incumbent. Basic research is missing.
- The analytic community lacks the institutional memory necessary for systematic coordination. Information exchange is limited.

1/See Dondero, in Zimmerman [115], "Concluding Remarks," pp. 1-2. The comment followed the presentation of D.W. Mader's paper--included in abridged form as appendix II.

- There is a need for an institutional focus which can provide the "memory" that is necessary to harmonize analysis and research efforts. 1/

These comments are apt. Our review suggests that there may be a number of analysts performing defense studies who are unaware of either the empirical validity of their tools, or the existing corpus of knowledge pertaining to their projects.

The point is that shortcomings--in "understanding war" and in determining good measures of effectiveness--constrain the usefulness of analysis in Defense Decision. 2/ We believe that improvements are possible, and that such improvements would be facilitated by establishing an institutional focus to "match the needs of analysis" with an institutional memory and a coordinated plan for research into the phenomenology of combat. 3/

1/The comment was not intended to diminish the important contributions of the Military Operations Research Society (MORS) and the Defense Documentation Center (DDC), but merely to point out a need beyond their scope. We address the role of this institutional focus later in the text. For now, let us note that much defense-related literature is not in DDC--e.g., many of the publications of the RAND Corporation and HERO; Allied studies and analyses from the SHAPE Technical Centre (STC), the Industrieanlagen-Betriebsgesellschaft (IABG), or the Defence Operational Analysis Establishment (DOAE); Ph.D. theses; and Operations Research articles. The institutional focus mentioned here would be required to remain abreast of relevant portions of all of this work in support of its advisory and referral functions--but it would not perform the repository services of the DDC.

2/Measures of effectiveness are inevitably approximate. "In Vietnam, to measure the progress of the war we are forced to use an avalance of statistical measures--incidents, defections, body counts, weapons lost and captured--all more or less unsatisfactory." See Quade and Boucher [74], pp. 361-362.

3/This is not to say that we advocate foreclosing on the individual Services' research programs. Indeed, pluralism seems to be called for to avoid premature and stifling closure on creativity in this area.

Because the accuracy with which Defense requirements are determined is a matter which affects highest national policy, transcends any individual Service, and involves international relationships; and because the development of "military effectiveness" data should be independent of the advocacy inherent in budgetary proceedings--it seems appropriate that the responsibility for providing this institutional focus should rest with the Nation's senior military advisers, the Joint Chiefs of Staff.

RECOMMENDATIONS TO THE SECRETARY OF DEFENSE

We recommend that the Secretary of Defense require the Chairman, Joint Chiefs of Staff to:

- Review current procedures for safeguarding and strengthening the empirical-theoretical foundation underlying the analytic representation of combat in Defense studies. This should include

  --identifying the empirical-theoretical limitations of such studies as are employed in high-level Defense decisionmaking; and

  --ascertaining the extent to which military and civilian defense analysts are aware of the empirical validity of their tools, and the existing corpus of knowledge pertaining to their projects.

And, as warranted by that review, to:

- Prepare plans and recommendations which would enable the Organization of the Joint Chiefs of Staff to serve as the Defense Establishment's principal analytic adviser on matters pertaining to the phenomenology of combat. Such plans and recommendations should consider the requirements for

  --developing adequate theories and empirical information about the integrative structures of combat which bind the Services together in a web of common concern and interests;

  --making recommendations to the Secretary of Defense on priorities for research, and on preferred theories and criteria to be used in defense analyses; and

  --establishing an institutional focus to provide the Defense analytic community with (1) a senior adviser, (2) a corporate memory, (3) a technical

forum, and (4) a center to disseminate information on the techniques of combat analysis.

AGENCY COMMENTS AND OUR EVALUATION

The Department of Defense advised us that

"OSD is in the process of establishing a formalized procedure for spot auditing of study efforts throughout DoD, concentrating on several mission areas each year. This overall review of the content, methodologies and use of studies and analysis will give close attention to modeling and other analytical tools employed."

The Department also commented on the "conditional" part of our recommendation, saying:

"The Studies, Analysis and Gaming Agency (SAGA) within the JCS presently performs analyses of alternative military strategies, force structures and postures in support of the responsibilities of the JCS to the Secretary of Defense. In this context they are the principal analytic advisers on matters pertaining to the 'phenomenology of combat'.

"The JCS already has a senior scientific and technical adviser within the SAGA organization. It is his responsibility to advise members of the JCS on issues of a scientific or technical nature. Since each DoD component has responsibilities for defense analyses that are not within the purview of the JCS, it would not be practical to establish an institutional focus for all Defense analysis within that organization."

The Department did not address the elements in this recommendation pertaining to:

- developing adequate theories and empirical information ***;
- making recommendations to the Secretary of Defense ***; or
- establishing an institutional focus to provide *** a corporate memory, a technical forum, and a center to disseminate information ***.

We believe the Department's plan to establish "a formalized procedure for spot auditing of study efforts throughout DoD" is an important initiative for safeguarding the analytic representation of combat in Defense studies.

We are less optimistic about the Department's response that it would not be practical to establish an institutional focus for all Defense analysis within the Organization of the Joint Chiefs of Staff. Indeed,

- If the Department's logic were applied to fiscal matters, it would be impractical for the ASD(Comptroller) to provide a consistent fiscal framework for between-Service decisions. In point of fact, there is a consistent fiscal framework to suport high-echelon Defense decisionmaking. We believe there is a similar need for a consistent analytical framework.

- Our recommendation is intended to provide a foundation for the centralized coordination of methodologies used in analyses supporting decisions on strategy, force planning, and inter-Service tradeoffs. It uses the term "adviser," and refers to "the integative structures of combat which bind the Services together," to preclude any interpretation of an intent to: (a) establish centralized control over all defense studies and analyses; or (b) infringe on the legitimate prerogatives of individual Services or Defense components. In short, the recommendation addresses matters on which the Services "recommend"--and the Secretary "decides."

- The Department has pointed out that "the Studies, Analysis and Gaming Agency (SAGA) within the JCS . . . are the principal analytic advisers on matters pertaining to the 'phenomenology of combat'." One would expect therefore, that their talents would be applied toward the creation of a consistent analytical framework for high-echelon Defense decisionmaking. Such action was not evident during our review, nor is it implied in the Department's response.

As noted in the Defense Resource Management Study:

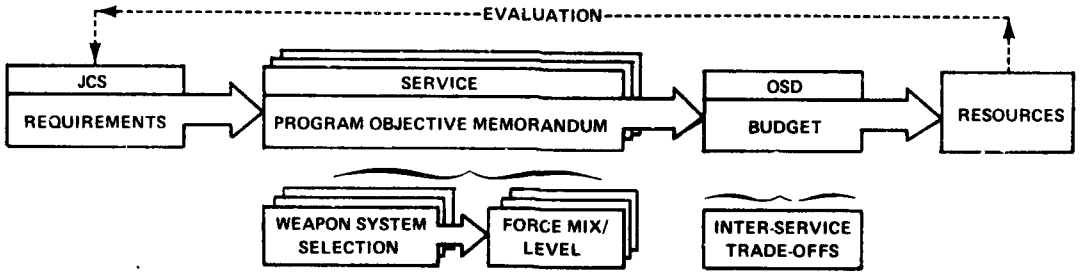
"The PPB System has never had an explicit measurement system for tracking the progress made in implementing approved programs. The heart of this matter

is the absence of objective performance standards. Program decisions are generally based on comparisons of estimated capabilities associated with alternative resource allocations. Analyses supporting such decision processes incorporate explicit management goals, scenarios, and support assumptions. Reporting systems that key on purchased manpower, equipment, or units (divisions, wings, or ships) are relatively meager reflections of the actual defense capabilities purchased. * * * Fiscal accounting, oriented to fiduciary responsibilities, does not provide adequate measures of program execution. Better feedback is needed, not only to monitor execution, but also to make adjustments to past decisions that, in turn, will motivate better execution." 1/

Textbook descriptions of the present PPB System (in the procedural sense) typically illustrate the process shown in figure 5.2. Several points are worth noting:

Figure 5.2

Evaluation Paradigm for the Modern Design for Defense Decision



1/See Rice, D. B., Defense Resource Management Study, Final Report, A Report Requested by the President and Submitted to the Secretary of Defense, Feb. 1979, p. 9.

- First, as noted in the Defense Resource Management Study (DRMS), the PPBS lacks an appropriate evaluation/feedback mechanism.

- Second, as Chapter 4 indicates, different analytical frameworks are used to support "Requirements," "POM," and "Budget" decisions. To repeat an earlier argument: "Analysis is unique to its framework. To go between frameworks is to bridge an analytic discontinuity involving far more than just mathematics ..." (see pp. 82-83). This is not an argument against examining various contingencies; it is simply a statement that the sacrifice of a consistent analytical framework on central issues may counterbalance any moderate advantage in variety. The potential impact of suboptimizations and methodological discontinuities in going from requirements to resources--and then to subsequent requirements--should be apparent. The importance of a consistent analytical framework to support Defense decisionmaking/evaluation should also be apparent.

- Third, while each Service and DoD component has evolved an analytic structure to support internal decisions, it is also desirable that those analyses be congruent with DoD objectives. The ASD (Comptroller) coordinates the Department's fiscal procedures; the coordination of analytic procedures is no less important.

By contrast, the Steadman report cited the JCS' inability to grapple with alternatives linked to resources when it recommended, among other things,

- "That the Secretary of Defense designate the Chairman, JCS as responsible for providing military advice from a national viewpoint on program and budget issues."

- "That the Chairman be given appropriate Joint Staff support to make broad program and budget judgments." 1/

This year's DRM Study reported similar weaknesses in the JCS' ability to grapple with alternatives linked to resources.

1/See Steadman [78], p. 69.

It concluded simply that "* * * more is needed if the JCS are to have a credible institutional role in the allocation of resources (the creation of capabilities) * * *." 1/

Our recommendations in this chapter are intended to strengthen what we believe is a continuing weakness in the design for Defense Decision. We believe they deserve the Department's further consideration.

1/See Rice [cited in fn 1, p. 107 of this report], p. 21.

CHAPTER 6

THE FOUNDATION FOR DEFENSE DECISION CAN BE IMPROVED-- FOR IT IS THE DECISIONMAKERS WHO CREATE THE INSTITUTIONAL PRESSURES FOR IMPROVEMENT

A major contention of this report has been that quantitative techniques have considerable potential as an aid in the analysis of defense issues, but that this potential is impaired by the "in-practice" design and management of quantitative tools. Improving these tools--providing better information for tomorrow's decisionmakers--has been the theme of this report and the purpose of our recommendations. The critical problems, in our view, are not technological, they are institutional. Their correction deserves the attention of all Defense decisionmakers.

MATTERS FOR CONSIDERATION BY DEFENSE DECISIONMAKERS

Warfare itself continues to include intuitive judgments that make it more an art than a science. So also Defense Decision, despite some growth in the scientific knowledge of war, must be recognized as drawing its strength and substance from military judgment. Quantitative methodology complements this judgement. It offers a framework that permits science and the expertise of diverse military specialties to be brought together--made explicit--and used to strengthen the Foundation for Defense Decision. This is its aim and opportunity.

In this report, we have focused on a very narrow segment of the DOD's studies and analysis program. We have attempted to instill just the right amount of skepticism in those who believe that quantitative methods can solve every problem--and just the right amount of belief in those who feel that it can solve none. That quantitative tools and methods have aided Defense Decision is without question. But if the findings of this and earlier studies are representative, the "in-principle" benefits of these tools have not been realized. For these powerful tools to achieve their potential, decisionmakers must create the "in-practice" pressures for improvement. This requires each decisionmaker to be

1. A knowledgeable consumer of quantitative information:

- The assumptions underlying a study's conclusions should serve as the basis for an open, explicit transfer of understanding to the decisionmaker.

- The conclusions are the user's, not the model's. If the analysis is not transparent, then the decisionmaker must base its credibility on the experience and institutional affiliation of the analyst.1/

2. An investor of knowledge:

- To the extent that a policy assisting model's assumptions are testable hypotheses, the question is--"What is being done to improve the foundation for tomorrow's decisions?" (see pp. 21-25).

- The creation of a National Training Center, supported by the Secretary of Defense, the Army, and the Defense Advanced Research Projects Agency, is a positive step in the effort to understand the phenomenology of combat (see pp. 93-96). Our recommendation for a principal analytic adviser (see pp. 104-105) is intended to further strengthen the empirical-theoretical basis for Defense Decision.

3. A concerned manager:

- Quantitative tools require management support and control if they are to provide reliable support for decisionmaking (see pp. 29-30).

- The VECTOR-2 User's Group and CCTC's Weapons Performance Data Base are important steps toward achieving a consistent framework for analysis. Our recommendations concerning configuration management and a "reference" data base (see p. 80) are intended to strengthen this aspect of Defense Decision.

Let us clearly understand that when we use quantitative tools for policy analysis, we are "playing with fire." Let us also recognize that if cavemen had not "played with fire,"

1/It is important to note that quantitative methodology merely complements the decisionmaker's judgment. There is no *a priori* guarantee that it will enable a novice to render expert judgments.

we would still be in the Stone Age. The tools have tremendous potential. The operational imperative is to manage and use them wisely.

Under a variety of names--such as operations research, computer modeling, and cost-effectiveness analysis--quantitative tools and methods have come to play a large and pervasive role in the analysis of public policy issues. That use of these methods can be fruitful and rewarding is without question. That they can be misused--with unfortunate choices and consequences--is also without question. What is unique about the present setting, however, is the extent to which practitioners of quantitative methodologies have become involved in Government (and military) information handling.

Weapon systems costing hundreds of millions or even billions of dollars, composition of future force mixes, and other defense planning and decisionmaking are often justified, in part, or supported by quantitative studies. Alternatively, the studies themselves have a cost. The Department of Defense formally requested $164.2 million for Studies and Analyses in FY 1980, while pointing out:

> "Studies are not conducted independent of other activities but are initiated by an office or a command in need of a study to help it reach a decision. Similarly, the total funding devoted to studies, which is substantial (about a quarter of a billion dollars), is expended in a great number of small sums."1/

This report has attempted to highlight the differences between earlier decisionmaking aids and today's quantitative tools. With traditional forms of decisionmaking, the difference between science and judgment was recognizable. In contrast, with today's complex policy problems, the decisionmaker's information is often provided by quantitative techniques which embody subtle admixtures of fact and opinion. Such techniques are inherent in mission budgeting, risk assessment, the evaluation of social programs, and the like.

1/Testimony of the Honorable William J. Perry before the Senate Appropriations Committee, Department of Defense Appropriations, Fiscal Year 1980, Part 4--Procurement/ R.D.T. & E., 96th Congress, First Session, pp. 169-170. The difficulties of determining DoD's modeling costs are described in fn 1, p. 45 of this report.

The theory from which these techniques are drawn assumes the existence of a well-defined problem. Because human judgment per se is not represented in the theory, it can be neglected in applications of the methodology. The neglect shows itself in a tendency to focus on the calculation and the results of calculation, and to ignore what is being calculated and why.

The analysis of a squishy problem requires judgments. To the extent that such judgments are not explicit, their relevance is questionable. To the extent that the responsible analyst/decisionmaker is unknown, the credibility of those judgments is suspect. And, to the extent that all are unknown--caveat emptor, let the buyer beware of anonymous opinions.

RECOMMENDATIONS TO THE CONGRESS

Congressional concern for the quality of quantitative decisionmaking tools is a matter of record. When reviewing quantitative studies or exercising its oversight authority, the Congress should require:

- An open, explicit understanding of the assumptions underlying a study's conclusions, or
- Knowledge of
  - -- the identity of the decisionmaker(s) involved in the study,
  - -- their background experience and institutional affiliation(s), and
  - -- the extent to which the model(s) used in the study have been appraised, possibly in the form of an explicit statement accompanying the results of the study.

In its oversight role, the Congress should also require agencies and departments to report on: (1) how current decisionmaking tools are being managed, and (2) what is being done to improve the tools and provide better answers to difficult public policy questions in the future. Numerous other questions could be applied to quantitative studies. For example:

- What alternatives were considered in the analysis? What criterion was used to compare the various alternatives? (Recall the difference in effectiveness produced when the antiaircraft guns of chapter 5 were analyzed according to their ability to

shoot down enemy aircraft as opposed to their ability to protect merchant ships. See pp. 97-98.)

- Answers to questions such as those proposed by J. A. Stockfisch might also be of interest. Is a number used in a quantitative study or analysis "the output of a model, or the result of some physical measurement? If it is the output of a model, *** has the model been validated by some independent test? If not the latter, then what is the structure of the model--i.e., what is the theory? If a model has been tested, or if a set of numbers are the result of physical testing or some other empirical source, then what was the experimental matrix and what are possible instrumentation errors, or what were the reporting methods employed? How was the data filtered and aggregated as it moved upwards (and often sideways) in the bureaucratic hierarchy? If the subjective assessments of individuals are used for certain kinds of data generation, who were these individuals and what has been their experience and institutional affiliation?"1/

And finally, in reviewing defense studies and analyses, the Congress should inquire as to how a particular program or weapon system contributes to the overall force level analyses. (Recall the example of suboptimization on p. 45.) In this way, the Congress can ascertain the links between the measures of effectiveness used in the acquisition process and those used at higher levels of analysis. Typical questions might include:

1. How are requirements ascertained?
2. How do they originate?
3. Who evaluates them?
4. How is the evaluation done?
5. How are competing interests resolved?

1/See Stockfisch [79], pp. vii-viii.

6. What conditions must exist before funds are committed (or requested) for the procurement of services, equipment, and other items which constitute military requirements?

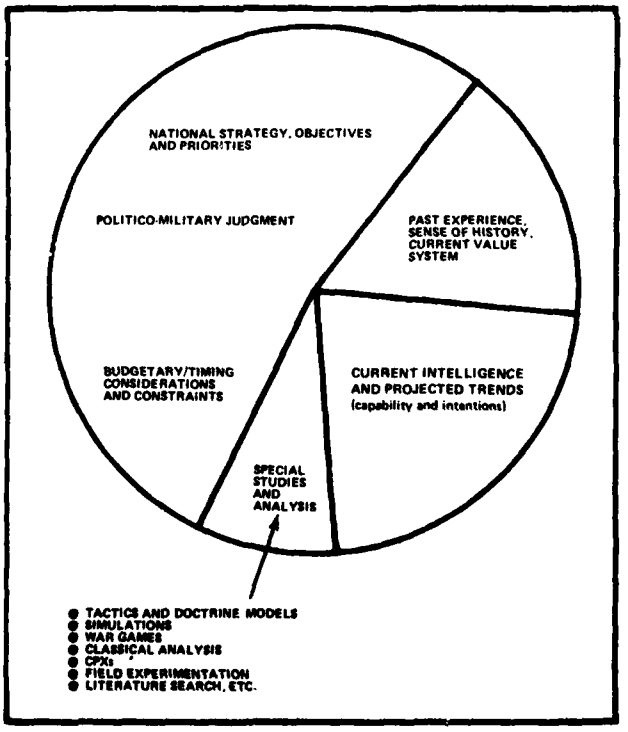
These are all matters which have considerable bearing on the combat readiness of the Armed Forces as well as on defense expenditures.

MODELS HAVE TO BE KEPT IN CONTEXT

[Dr. F. B. Kapper is the former Scientific and Technical Adviser for the Organization, Joint Chiefs of Staff. This Appendix presents his remarks, abridged from "Session I," THEATER-LEVEL GAMING AND ANALYSIS WORKSHOP FOR FORCE PLANNING, Volume I-PROCEEDINGS, Office of Naval Research, 1979.]

Models have to be kept in context and what they produce has to be kept in context as well (Slide 5-1). The decision maker uses computer results as input to his decision making process, not as his basis for a decision. He also uses quite a few other things. He uses the capability and intent that he gets from the intelligence side of the house, and he uses his own knowledge of military history. I think if you look at what models produce, unless you keep this in perspective, you're going to lose track of a model's relativity.

SLIDE 5-1 KEY FACTORS INFLUENCING THE POLICY/DECISION MAKER
(A Subjective Viewpoint)



Models are a means to an end; models are tools. They are not the end objective in themselves, except to the model maker and the model designer at a particular stage in the model's development. From a user's standpoint,... they are tools, and very helpful ones. But I think that models have got to be put in the more meaningful context of studies and analyses. We don't just run a model or models--we do it to study something or to analyze something, some kind of phenomena. So, again, it is important to keep in mind that models are an input to studies and analyses.

Studies and analyses have many purposes, but their main purpose is to understand the phenomena--that's really the most important thing. If you can predict, you can control... but the primary purpose of models is to gain an understanding of very complex phenomena.

Now, in talking about where the JCS uses models relative to types of decisions, in a theater context, we do use them relative to the support of the Joint Strategic Objectives Plan (JSOP), for the Joint Force Memorandum (JFM) and in special analyses. One of the reasons why we use them, of course, is to determine the relative risk for particular force postures. We don't determine risk in a vacuum; it's relative to a threat. So, I think that's one of the most important things vis-a-vis model use in the JCS.

In the JCS we also do a great many special analyses in which we use theater combat models. The mutual balanced force reduction (MBFR) series of analyses use theater combat models. The Short Pact Attack Study, or scenarios, investigated the ability, and the desirability if you will, of the Pact to initiate attacks given certain short duration preparation times. We use models a great deal in the CAP series of studies for logistics and mobility... In addition, we used models in the Military Committee Special Study Group (MCSSG) to look at the relative force balance between NATO and the Warsaw Pact. So, in terms of the types of decisions, we do talk about operation plans a little, and we talk about the Joint Strategic Planning System publications and their support. We do these studies to identify shortfalls and weaknesses in plans, programs, force structures, and so on.

Now, let me make another point. It's very, very difficult to say my study or my analysis really resulted in this particular policy being formulated, or it resulted in that program being initiated or cancelled. The reason why it's tough to do is very simple. If the decisionmaker is a good decisionmaker, he takes that study and analysis and uses it as input. If he didn't he wouldn't be earning his keep.

Sometimes the analyst can identify where his work did save money, or where it did help some things. An example is the shelter program study which we did about two or three years ago. It was very crucial in the overall orientation for NATO to go to aircraft shelters. Part of the data, incidentally, came from the Israeli/Arab war.

Now, what are some of the problems? Let's focus on three--time, cost, and value. That's kind of a general overhead, but these things really impinge on people who are in the studies and analyses business. And, there are several aspects to time, cost, and value.

For example, it takes a long time to train somebody with respect to using a particular model. There are ways in which to shorten that time, but it is a key problem. We are constantly bringing people in and out of the studies and analyses business. Basically they need to be trained how to use a model more so after they learn to punch this button, to fill in the data this way, and so forth. What they need to know are its weaknesses, where it is appropriately used and not used, and what are the complex aspects of the model that as users they should know. One bright chap used video-tape to document some key aspects of his algorithm--that's a great idea. And, there are other ways in which to get over the lead time.

Models are frequently designed and developed that ignore the availability of existing data bases, or ignore whether or not data are available for the particular action they want to create or simulate. There are data base weaknesses, and I don't think we can talk about models without at least raising the data base weakness issue. And this of course relates to the crucial issues of standardization and validation.

Comparability of results from different models is something that gives us all some concern. Assuming that we all start with the same assumptions, as we pretty much did in the trilateral talks, it is still possible to have vastly different results. We still had differences between the FRG, the UK and the US, and then the SHAPE Technical Center came in and harmonized everything for us which was a big help.

The phenomenology of combat, as far as I'm concerned, is not as well understood as it needs to be. In all honesty, I don't think we fully understand the interaction of combined arms. Take a typical ground force situation, with one guy looking at it. Now, someone introduces all kinds of air/ground interaction and then it gets somewhat complex. Then someone puts in some tactical nuclear or chemical munitions

and that really creates a lot of complexities that I really don't think we understand. I think we should try to look at the basic phenomena and try to get a better handle on the essentials. I don't think we do enough in this regard.

Also, I have a feeling that we often use model studies and analyses to make up for quantitative weakness, vis-a-vis NATO/Warsaw Pact. I think we can play qualitative games in this area only so long before there will be some severe prices to pay.

Lastly, there's a lack of frequency in the communication that we have with our allies, particularly our German and British allies. I think we really need to pull our allies into our modeling business, into our studies, analyses, and gaming business, if you will. They have learned a lot of lessons and have perspectives which can be of great benefit to us all.

ORIGIN AND DEVELOPMENT OF RATES OF ADVANCE (1954 - 1969)

[Abridged from "Chapter 2 - Historical Origins," NATO COMBAT CAPABILITIES STUDY, Volume VI - Rates of Advance of Theater Forces, Research Analysis Corporation, RAC-CR-56, June 1972.]

In 1954 CORG 1/ published two papers, one by Colonel A. D. Hulse and the other by Lt. Colonel N. W. Parsons, 2/ containing the rates-of-advance tables which are the foundation for those in use today. Since they are the seminal documents, they will be rather extensively quoted.

Colonel Hulse's paper contains an analysis of twenty historical examples from WWII (three from Italy, the remainder from France, Belgium, and Germany) of advances made by U.S. armored divisions or mixed task forces within those divisions. Concerning these examples, Hulse wrote,

> "It is worthy of note that deep armored drives in enemy rear areas with scattered, disorganized resistance netted on the average of 25-30 miles per day. There are examples of individual daily gains of up to ninety miles. In these cases, however, all resistance was by-passed, no containing forces were left in the rear, and the leading elements were light, highly mobile reconnaissance units.... In all of these historical examples, it should be remembered, the advances were made with air support on call, column cover, and/or preceeded by heavy air strikes. Interference from enemy air was insignificant."
>
> "U.S. ground operations in Europe during WWII were in general unaffected by the tiny German air effort. Tactical troop movements and

---

1/CORG, the Combat Operations Research Group, was at that time a field activity of the Johns Hopkins University Operations Research Office (ORO) which was the predecessor of the Research Analysis Corporation (RAC).

2/GAO note. For a reproduction and cogent analysis of the papers, see: Wainstein [106].

> logistics operations were carried on day and night. The movement capability behind the line of contact therefore approached a theoretical 100%. On the other hand, German troop movements and resupply operations were in general limited to hours of darkness. Further, the road and rail net capacity in the Germans' rear was greatly reduced by unceasing Allied interdiction operations. The 'movement capability' then of the enemy was on the order of 30%, if we consider that it might have been 100% with complete absence of Allied air."

In duration, the examples cited range from a few hours to 13 days. Rates of advance are given in miles per day, miles per hour, and even (against very heavy resistance) in yards per hour. Translation from one measurement basis to another is not consistent. For example, one 2-day battle is noted as involving 11 hours of actual attack while another 2-day battle involved 28 hours of attack. Twelve hours per day and 14 hours per day are also used. These inconsistencies were later obscured with a consequent distortion of the data.

LTC Norman W. Parsons analyzed 32 examples of battalion sized infantry unit attacks. Three were in Guadalcanal, seven in Italy, and the remainder in Western Europe. Concerning his sources of data, Colonel Parsons wrote,

> "It should be noted that all of the examples are from narratives in which the time of attack and the time that the objective was reached are stated and the location of the line of departure and of the objective are shown on maps. The necessity of knowing each of these facts has severely limited the number of examples which could be considered and tabulated. However, in many other examples in which the lack of data did not permit calculation of exact rates of advance, it was noted that the rates of movement were of the same magnitude as those tabulated. It should be noted that there were many unsuccessful attacks by the same units in the same general terrain and against the same general types of enemy resistance. Those unsuccessful attacks have not been tabulated.... Among the examples studied it was found that the rates of advance varied from a maximum of 1760 yards per hour to a minimum of 114 yards per hour...

the mean rate of advance is 563 yards per hour. Seventy-five percent of the examples have a rate of advance of 890 yards per hour (one-half mile per hour) or less."

The fact that Hulse and Parsons based their results on successful armored and infantry unit attacks was also lost in subsequent modifications of their work. Another characteristic of their data, that they applied to infantry battalions and to armored battalions and combat commands, also faded with the passage of time.

* * * * *

As Hulse and Parsons readily recognized, there were many significant factors affecting rates of advance that they could not take into account either because they were intangible or because they were transient conditions that had not been adequately described in the source documents. In the end they produced tables of ranges of rates of advance that took into account only three things: whether the attacking force was composed primarily of infantry or of armored units, a judgment of the degree to which the defender resisted the attacking force, and a gross evaluation of the essentially permanent characteristics of the terrain involved.

A careful examination of the way in which Hulse and Parson described these three factors is important to an understanding of subsequent events. The first point to note is that the basic data applied to infantry battalions and to armored battalions and combat commands. These fairly specific unit-size designations were later relabelled as infantry or armored units, then simply as infantry or armor with no size implications whatsoever, and finally they became infantry or armored divisions.

The second point to note is that Hulse and Parsons used words as their primary means of describing the amount or degree of enemy resistance. They used five categories: none, light, moderate, heavy, and very heavy. Only parenthetically did they relate these verbal descriptions to what they called "rough and approximate" numerical force ratios. They identified light resistance as associated with situations in which the attacker to defender force ratio was 5 or 6 or more to 1; moderate resistance they associated with a force ratio of 4 to 1; heavy with 3 to 1; and very heavy resistance implied force ratios of 2 to 1 or less. As we shall see, this essentially simple and understandable verbal scale was later inverted, torn apart, and stretched by other rates-of-advance table developers.

The third point to note is that Hulse and Parsons used five terrain types which they defined as follows:

1. Open country; lightly wooded, slightly rolling, elevation changes less than 30 meters per mile and not more than 25 percent woods cover.

2. Moderately open country; wooded and rolling, elevation changes 30 to 50 meters per mile or 25 percent to 50 percent woods cover.

3. Moderately close country; wooded and hilly, elevation changes 50 to 100 meters per mile or 50 percent to 75 percent woods cover.

4. Close country; heavily wooded or sharp close hills, limited visibility, elevation changes 100 to 300 meters per mile or over 75 percent woods cover.

5. Mountainous; elevation changes over 300 meters per mile.

These five types of terrain were soon reduced to three with slightly different definitions. Three became six with introduction of the notion that a militarily significant barrier could be present in any of the three recognized types of terrain. And the most recent change has been to go back to five types of terrain with a sixth "terrain type" which is only a barrier.

The final point to note is that Hulse and Parsons gave a range of rates of advance for each combination of terrain type and degree of enemy resistance. (See tables 1 and 2.) Their tables contain 40 ranges of rates of advance against resistance plus 10 against no resistance. Later versions went from as low as 30 ranges to as high as 840 specific rates of advance. (It should be noted that many of those specific rates were zero.) The maximum rate has ranged from Hulse and Parsons' 30 miles (48km) per day down to about 15 or 16 miles (24km) per day and up to 62 miles (100km) per day.

GAO NOTE: Wainstein's 1973 examination of the Parsons and Hulse papers [106] led him to conclude that

-- the papers were rough, hasty efforts, not comprehensive pieces of research adequate for broad applicability; and

-- both papers were utilized far beyond the obviously limited intent of the authors.

Now that we have described the starting conditions, let us trace through the modifications and alterations that will bring us to the rates-of-advance tables in use today.

* * * * *

Table 1

ARMORED BATTALIONS AND COMBAT COMMANDS
Proposed Rates of Advance
(For use in map maneuvers)

	Degree of enemy resistance (and rough force ratios)				
Type of terrain	None	Light and scattered (6-1 or more)	Moderate 4-1	Heavy 3-1	Very heavy 2-1 or less
Open	7-10 (mph)	25-30 (mpd)	15-20 (mpd)	6-12 (mpd)	500-700 (yds/hr)
Moderately open	5-7 (mph)	20-25 (mpd)	6-15 (mpd)	4-8 (mpd)	300-500 (yds/hr)
Moderately close	3-4 (mph)	15-20 (mpd)	5-12 (mpd)	200-500 (yds/hr)	100-200 (yds/hr)
Close	1-2 (mph)	7-10 (mpd)	200-300 (yds/hr)	100-200 (yds/hr)	0-100 (yds/hr)
Mountainous	0-1 (mph)	3-6 (mpd)	--	--	--

Table 2

RATES OF ADVANCE OF INFANTRY BATTALIONS
(In yards per hour)

	Amount of enemy resistance (and approximate force ratios)				
Type of terrain	None	Light FR 5-1 or more	Moderate FR 4-1	Heavy FR 3-1	Very heavy FR 2-1 or less
Open	1400-2500	800-1200	500-800	400-600	300-500
Moderately open	1000-2000	600-1000	400-600	300-500	200-400
Moderately close	900-1500	500-900	200-500	150-350	100-300
Close	800-1200	400-700	200-400	150-250	100-200
Mountainous	500-800	300-500	200-300	150-250	100-200

1957 - CONARC WAR GAMING HANDBOOK

The War Gaming Handbook published in 1957 by the United States Continental Army Command (CONARC) contained the same tables and credited the original sources. However, CONARC made some additions to the table for armored battalions and combat commands. In every instance where the original table gave a rate in miles per day, CONARC parenthetically added rates in yards per hour based on a 14-hour day. This modification presented all of the rates in both tables on a per-hour basis.

The original designation on the abscissa was "Degree of Enemy Resistance." CONARC changed that to "Force Ratio," made the original "rough and approximate" force ratios the primary column headings, and relegated to parentheses the verbal descriptions which had been primary. They also added letter designation A through E to identify the terrain conditions "Open" through "Mountainous."

1958 - CONARC SPECIAL WAR GAMING HANDBOOK

In this handbook CONARC made a number of changes to the rates-of-advance tables. They dropped terrain types B and D and the numbers associated with them. The remaining terrain types--open, moderately close, and mountainous--were relabeled as types A, B, and C, and they were given expanded definitions. Essentially type A was very favorable for armored operations, type B was marginal, and type C was "in no way suitable for armored operations." Also dropped were the miles-per-day rates that had been in the original tables and only the yards-per-hour rates that had been derived the preceding year on the basis of a 14-hour day remained. The fact that these rates had been so derived was not mentioned. CONARC also removed battalions and combat commands from the table headings and used "Units" in their place.

* * * * *

At this point the tables contained 30 separate ranges of movement rates--15 for infantry units and 15 for armored units in three different kinds of terrain against 5 degrees of enemy resistance.

1959 - 1960 ANALYTICAL WAR GAME CONTROL MANUAL

No sources whatsoever for the rates-of-advance tables in this manual, published by the Army War College (AWC), were specifically cited. Both the format and the contents bore little resemblance to tables in earlier handbooks. However, it can be clearly shown that the AWC tables of rates of

advance are directly derived from those in the CONARC Special War Gaming Handbook.

Earlier tables treated degree of enemy resistance and force ratio as essentially synonymous and interchangeable ways of describing a single characteristic of a combat engagement. The War College separated them. Their tables had six degrees of enemy resistance, described in terms of the defender's posture, across the abscissa, and ten numerical force ratios along the ordinate. The postures given were:

Defender Fortified Zone

Defender Prepared Position

Meeting Engagement

Defender Delaying Action

Defender Retiring

Defender Disorganized Retreat

and the force ratios were neatly graduated from equal to or less than one-half to one to equal to or greater than five to one.

Definitions for terrain types A, B, and C were essentially identical with those in the CONARC Manual, but the War College introduced the concept of "barriers," without explanation. Since barriers could be present in any type of terrain, they effectively doubled the number of tables. There were six terrain classifications (three without and three with barriers), six postures, 10 force ratios, and two kinds of attacking forces (infantry and armored). The War College therefore required 720 entries to fill out their rates-of-advance tables. Each entry was a specific rate, not a range of probable rates.

To go from the 30 ranges of rates in the CONARC Handbook to the 720 specific rates in the AWC Manual, the War College set up seven equivalences to define the limiting rates in their tables in accordance with the limiting rates in the earlier tables and then filled in between the limits on the basis of essentially linear interpolation.

* * * * *

We have noted earlier that CONARC substituted "units" for the original identifications of battalions and combat

commands. The War College went one step further and dropped "units"; their tables were simply identified as containing "armored" and "infantry" rates of advance.

1963 - RAC TP-76, QUICK GAME

The next step is to the Quick Gaming Manual published by RAC in 1963. A direct connection is readily established because the AWC Manual is cited as the source of the basic data. However, those basic data were much modified.

* * * * *

There were numerous differences. A new posture had been added. Data for this posture, "Defense of a Hastily Prepared Position," as well as adjustment of the data for the Meeting Engagement posture were "derived by graphical interpolation between the values for 'Prepared Position' and 'Defender Delaying' assuming equal spacing along the abscissa for each posture." All the postures were defined, which had not been done before.

Where the War College Manual gave rates of advance for force ratios of one-half or less and two-thirds, the Quick Game Manual gave no rates for force ratios of less than one.

But the major difference was that the source rates, which were in terms of yards per hour for battalions, had now been transformed into miles per day for divisions! Unfortunately, the details of the transformation were not recorded, only the results and some general statements about what was done. The process included at least one error. The author wrote that: "The basic data are in [the 1959-1960] Army War College Control Manual. These data consist of hourly rates of advance for infantry and armor divisions." The data actually applied to battalions and armored combat commands, but that knowledge had been lost and the AWC manual failed to give any information about the size of units to which their data applied.

The author of the Quick Game Manual went on to say:

> "Quick Gaming requires aggregation to a daily (24 hour) rate. At first glance one might assume that the daily rate might be the yards per hour . . . times 24; or, one might assume a 'typical day' of, say, 10 hours. Neither, of course, is correct. . . . A detailed analysis was available for Korea (1950 - 1953).

GAO NOTE: In a 1971 report, 1/Dondero, et al., criticized the adequacy of the Korean analysis in that:

- only two campaigns were included,
- no consideration was made of the flow of reserves and supplies over the road net, and
- all the forces involved were not included for the reason that they were either committed for a short period of time or they were not large enough to have been located on a day-to-day basis.

These data were used to develop daily rates of advance for the representative tactical postures and force ratios. They become part of the control manual for exercise JIGSAW, September 1955, in the Far East Theater. (The basic JIGSAW documents are unavailable. The pertinent data were preserved for use here in unpublished notes of J. W. Johnson, RAC.)

"The hourly data . . . [from the War College manual] and the daily data. . . [from Exercise JIGSAW], given the assumption that type B terrain is representative of Korea, can be combined to produce for each type of combat the average number of hours per day during which movement would occur. These numbers of hours are shown in [see Table 3]."

1/See Dondero, et al., [25].

Table 3

Hours of Intense Combat per Day

(During Which Movement Occurs)

FORCE RATIO	FORTIFIED ZONE[a]	PREPARED POSITION	HASTY POSITION[b]	MEETING ENGAGEMENTS[c]	DEFENDER DELAYING	DEFENDER RETIRING	ROUTE
1.0	0.5	1.6	4.0	5.2	7.9	9.5	10.2
1.5	0.8	2.0	4.8	7.2	9.0	9.6	13.3
2.0	1.4	3.0	5.7	8.0	10.1	11.8	16.0
2.5	2.5	4.4	6.9	8.8	10.9	12.8	18.2
3.0	3.7	5.8	7.7	9.4	11.5	13.8	20.1
3.5	4.8	6.7	8.2	10.0	11.9	14.6	21.3
4.0	5.6	7.2	8.6	10.4	12.1	15.4	22.4
4.5	6.1	7.6	9.2	10.8	12.5	16.0	22.7
5.0	6.6	8.0	9.7	11.4	13.2	16.4	23.0

[a]This column derived from extrapolation from the other postures.

[b]This column derived by graphical interpolation between Prepared Position and Delay.

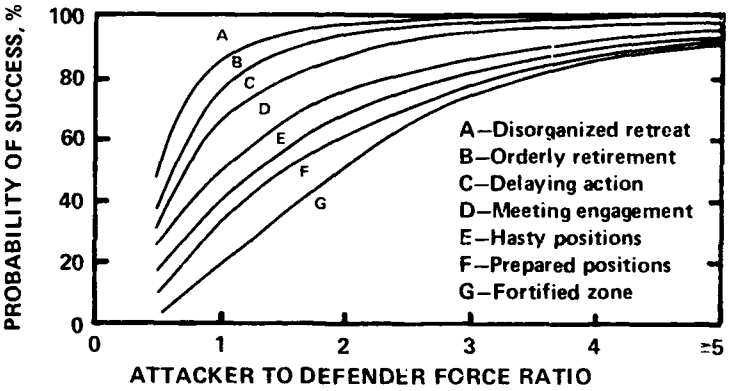
[c]This column adjusted by graphical interpolation between Prepared Position and Delay.

It is difficult to comment on Table 3 except to note that its major characteristic, which can be determined by inspection, is that it presumes armed combat to be characterized by consistent cause-effect relationships and an orderliness that is not evident in historical records.

> "Given the yards per hour in the . . . [War College manual] and the number of hours in that table, the miles per day were computed. There was one additional consideration at this point: the resulting rates of advance were obviously high for the center of mass of a division under 'Route' or 'Disorganized Retreat'. This discrepancy was eliminated by graphical interpolation after appeal to historical norms as given in ORO TP-10 [a paper by Marshall Andrews titled "Rates of Advance in Land Attack Against Unprepared Forces"].
>
> "The final step in constructing this table of rates of probable advance was to take the rates of advance if the attack was successful and apply the probabilities of success given . . . [in Figure 1]. This produced the 'expected value' of the advance per day for the attacker."

Figure 1

Probability of Success Curves



This final step introduced a major transformation in the nature of rates-of-advance tables. Up to this point they had been applicable only in those cases where, by other means, it had been decided that the attacking force was successful. Warfare had been modeled as a series of individual engagements in which the attacking force would advance if it were successful and would not advance if it were unsuccessful. In Quick Gaming, because probability of success, from Figure 1, was built into the rates of advance, every engagement is an average engagement. The tables purport to show the average advance that would be expected for every defined combination of force ratio, defender posture, and terrain type.

* * * * *

1969 - FOREWON

Computerized Quick Game was renamed ATLAS; ATLAS became the battle model component of the FOREWON System; and FOREWON was used in the preparation of the current ASOP. Thus, we have followed rates of advance from their origin in 1954 through to today.

BIBLIOGRAPHY 1/

[1] Aldrich, J. R. and J. R. Bode, The Aggregation Problem in Models of Theater-Level Warfare, BDM/W-77-129-BR, The BDM Corporation, Vienna, Virginia, Sept. 1977.

[2] Anderson, L. B., et al., IDA Ground-Air Model I (IDAGAM I), R-199, Volume 1: Comprehensive Description (AD A017 202); Volume 2: Definitions of Variables (AD A004 539); Volume 3: Detailed Descriptions of Selected Portions (AD A004 540); Volume 4: Documentation; Volume 5: Testing (U) (AD C000 941L) CONFIDENTIAL, Institute for Defense Analyses, Arlington, Virginia, Oct. 1974.

[3] _____, Modifications to IDAGAM I, P-1034 (AD B001 980L), Institute for Defense Analyses, Arlington, Virginia, Oct. 1974.

[4] Anderson, L. B. and R. E. Douglas, A Data Base for IDAGAM I (U), P-1033 (AD C000 940L) SECRET, Institute for Defense Analyses, Arlington, Virginia, Oct. 1974.

[5] Best, R. J., A Study of Battle Casualties Among Equivalent Opposing Forces: Korea, September 1950, ORO-T-23(FEC) (AD 2885), Operations Research Office, The Johns Hopkins University, Chevy Chase, Maryland, Apr. 1952.

[6] _____, Casualties and the Dynamics of Combat (U), RAC-TP-185 (AD 372 260) CONFIDENTIAL, Research Analysis Corporation, McLean, Virginia, Mar. 1966.

[7] Bode, J. R., Indices of Effectiveness in General Purpose Force Analysis, BDM/W-74-070-TR, The BDM Corporation, Vienna, Virginia, Oct. 1974.

[8] Bonder, S., "Systems Analysis--A Purely Intellectual Activity", Military Review, Volume 51, Number 2, 1971, pp. 14-23.

1/ Whenever possible, we have included, in parentheses, the "AD" document numbers used by the Defense Documentation Center. Requests for these documents may be addressed to the Defense Documentation Center, Defense Logistics Agency, Cameron Station, Alexandria, Virginia 22314.

[9] _____, "An Overview of Land Battle Modeling in the U.S.," Proceedings of the 13th U.S. Army Operations Research Symposium, Volume I, (AD A007 126), Washington, D. C., Nov. 1974, pp. 73-88.

[10] _____, "Changing the Future of OR," Plenary Paper Presented at the IFORS 8th Triennial Conference on Operations Research, Jun. 22, 1978.

[11] Bracken, J., et al., Methodologies for General Purpose Forces Planning, Volume II: Ground-Air Campaign Model (GACAM) (U), R-175 SECRET, Institute for Defense Analyses, Arlington, Virginia, Mar. 1971.

[12] Brewer, G. D., "What Ever Happened to Professionalism?", Interfaces, Volume 8, Number 4, Aug. 1978, pp. 63-72.

[13] Brewer, G. D. and O. Hall, Policy Analysis by Computer Simulation: The Need for Appraisal, P-4893, The RAND Corporation, Santa Monica, California, Aug. 1972.

[14] Busse, J. J., "An Attempt to Verify Lanchester's Equations," Paper 9.1, Developments in Operations Research, Volume 2, B. Avi-Itzhak, et al., ed., Gordon and Breach Science Publishers, New York, 1971, pp. 587-597.

[15] CEM--For model documentation, refer to: [61].

[16] Clark, D. K., Casualties As a Measure of the Loss of Combat Effectiveness of an Infantry Battalion, ORO-T-289 (AD 59384), Operations Research Office, The Johns Hopkins University, Chevy Chase, Maryland, Aug. 1954.

[17] Command and Control Technical Center, Institute for Defense Analyses Ground-Air Model I (IDAGAM I): User's Manual, CSM UM 201-75 (AD B007 074L), Defense Communications Agency; Washington, D.C., May 1975.

[18] _____, Institute for Defense Analyses Ground-Air Model I (IDAGAM I), (Description of Coding and Logic Changes Implemented for IDAGAM II), CSM MM 201-75 (Supplement I), (AD B017 179), Defense Communications Agency, Washington, D.C., Oct. 1976.

[19] _____, Institute for Defense Analyses Ground-Air Model II (IDAGAM II), User's Manual, CSM UM 201-75 (AD B019 697), Defense Communications Agency, Washington, D.C., Oct. 1976.

[20] _____, Weapon Data Base Design and Documentation, TR-122-77, Defense Communications Agency, Washington, D.C., Sept. 1977.

[21] _____, Derivation of Selected IDAGAM II Input Values for 1980 Level Forces (U), TM 161-77 (AD C013 380L) SECRET, Defense Communications Agency, Washington, D.C., Nov. 1977.

[22] _____, VECTOR-2 System for Simulation of Theater-Level Combat, User's Manual, Part I and Part II, CSM UM 244-78, Feb. 1978; Maintenance Manual, CSM MM 244-79, Jan. 1979; Technical Memorandum, TM 201-79, Jan. 1979; Defense Communications Agency, Washington, D.C.

[23] Cordesman, A. H., Materials for the C-5 Working Group "Gaming and Simulation" of the 35th Session of the Military Operations Research Society, (AD A039 960) Office of the Assistant Secretary of Defense (Intelligence), 1974.

[24] Dondero, L. J., et al., NATO Combat Capabilities Study, Volume VI-Rates of Advance of Theater Forces, RAC-CR-56 (AD 901 744), Research Analysis Corporation, McLean, Virginia, Jun. 1972.

[25] Dondero, L. J., et al., Methodology for Force Requirements Determination (MEFORD) (U), RAC-R-121 (AD 515 716L) SECRET, Research Analysis Corporation, McLean, Virginia, May 1971.

[26] Dudzinsky, S. J. and J. Digby, The Strategic and Tactical Implications of New Weapon Technologies, P-5765 (AD A040 867), The RAND Corporation, Santa Monica, California, Dec. 1976.

[27] Dupuy, T. N., et al., Historical Trends Related to Weapon Lethality (AD 458 760), Historical Evaluation and Research Organization, Dunn Loring, Virginia, Oct. 1964.

[28] _____, Historical Trends Related to Weapon Lethality, Annex Volume III, Comparative Analysis of Historical Studies, (AD 458 759), Historical Evaluation and Research Organization, Dunn Loring, Virginia, Oct. 1964.

[29] _____, Opposed Rates of Advance of Large Forces in Europe (ORALFORE), (AD 902 830L), Historical Evaluation and Research Organization, Dunn Loring, Virginia , Aug. 1972.

[30] Engel, J. H., "A Verification of Lanchester's Law," Operations Research 2, 1954, pp. 163-171.

[31] Fain, J. B., "The Lanchester Equations and Historical Warfare: An Analysis of Sixty World War II Land Engagements," History, Numbers, and War, Volume 1, Number 1, Spring 1977, pp. 34-52.

[32] Fain, W. W. and J. B. Fain, Validation of Combat Models Against Historical Data, Professional Paper No. 27 (AD 704 744), The Center for Naval Analyses, Arlington, Virginia, Apr. 1970.

[33] Farrell, R. L. and R. J. Freedman, Investigations of the Variation of Combat Model Predictions with Terrain Line of Sight, Report AMSAA-1 FR75-1, Vector Research Incorporated, Ann Arbor, Michigan, Jan. 1975.

[34] Gary, et al., Federally Supported Mathematical Models: Survey and Analysis, RANN, The National Science Foundation, Washington, D.C., 1975.

[35] Gass, S. I., ed., Utility and Use of Large Scale Mathematical Models, Proceedings: April 28-29, 1977 Workshop, U.S. Department of Commerce, Washington, D.C., May 1979.

[36] Goad, R., Predictive Equations for Opposed Movement and Casualty Rates for Land Forces, STC PP-90 (AD 921 726), SHAPE Technical Centre, The Hague, Netherlands, Aug. 1974.

[37] Graves, H. K., "FEBA Movement and Attrition Process," Final Proceedings of the Fourteenth Annual U.S. Army Operations Research Symposiumm, Volume II, (AD B009 956L), U.S. Army Materiel Development and Readiness Command, Alexandria, Virginia, 1976, pp. 1053-1066.

[38] Greenberger, M. and R. Richels, "Assessing Energy Policy Models: Current State and Future Directions," final draft of paper for the Annual Review of Energy, Volume 4, Apr. 1979.

[39] Grosse, R. N., "Principles of Cost Effectiveness Analysis," Topics in Military Operations Research, The University of Michigan Engineering Summer Conferences, Jul. 25-29, 1966.

[40] Healy, J. G. and A. J. Rolfe, Testing of IDAGAM I (U), P-1032 (AD C000 942L) CONFIDENTIAL, Institute for Defense Analyses, Arlington, Virginia, Oct. 1974.

[41] Helmbold, R. L., Lanchester Parameters for Some Battles of the Last Two Hundred Years, CORG-SP-122 (AD 481 201), Combat Operations Research Group, Fort Belvoir, Virginia, 1961.

[42] _____, Historical Data and Lanchester's Theory of Combat, CORG-SP-128 (AD 480 975), Combat Operations Research Group, Fort Belvoir, Virginia, Jul. 1961.

[43] _____, Historical Data and Lanchester's Theory of Combat: Part II, CORG-SP-190 (AD 480 109), Combat Operations Research Group, Fort Belvoir, Virginia, Aug. 1964.

[44] _____, "Some Observations on the Use of Lanchester's Theory for Prediction," Operations Research 12, 1964, pp. 778-781.

[45] _____, Probability of Victory in Land Combat As Related to Force Ratio, P-4119 (AD 696 489), The RAND Corporation, Santa Monica, California, Oct. 1969.

[46] _____, Air Battles and Ground Battles--A Common Pattern?, P-4584, The RAND Corporation, Santa Monica, California, Jan. 1971

[47] _____, Decision in Battle: Breakpoint Hypotheses and Engagement Termination Data, R-772-PR, The RAND Corporation, Santa Monica, California, Jun. 1971.

[48] Hitch, C. J., "Plans, Programs, and Budgets in the Department of Defense," Operations Research 11, 1963, pp. 1-17.

[49] Hitch, C. J. and R. N. McKean, The Economics of Defense in the Nuclear Age, Anthenum Press, New York, 1965 (originally published by Harvard University Press, 1960).

[50] Honig, J., chm., Review of Selected Army Models, (AD 887 175), Models Review Committee, Department of the Army, Washington, D.C., May 1971.

[51] IDAGAM--For model documentation, refer to: [2]-[4], [17]-[19], [21], and [40].

[52] Karr, A. F., On the CONAF Evaluation Model, P-1210 (AD A030 232), Institute for Defense Analyses, Arlington, Virginia, Aug. 1976.

[53] _____, On the Lulejian-I Combat Model, P-1182 (AD A030 171), Institute for Defense Analyses, Arlington, Virginia, Aug. 1976.

[54] _____, On the IDAGAM I Combat Model, P-1265 (AD A043 530), Institute for Defense Analyses, Arlington, Virginia, Jun. 1977.

[55] _____, Additional Remarks on Attrition and FEBA Movement Computations in the Lulejian-I Combat Model, P-1277 (AD A047 835), Institute for Defense Analyses, Arlington, Virginia, Nov. 1977.

[56] _____, Review and Critique of the VECTOR-2 Combat Model, P-1315, Institute for Defense Analyses, Arlington, Virginia, Dec. 1977.

[57] Kerlin, E. P., NATO Combat Capabilities Study, Volume IV - Impact of Losses and Replacements on Unit Combat Capability, RAC-CR-56 (AD 901 743), Research Analysis Corporation, McLean, Virginia, Jun. 1972.

[58] Kerlin, E. P. and R. H. Cole, ATLAS: A Tactical, Logistical, and Air Simulation: Documentation and User's Guide, RAC-TP-338 (AD 850 355), Research Analysis Corporation, McLean, Virginia, Apr.1969.

[59] Kerlin, E. P., et al., Computerized QUICK GAME: A Theater-Level Combat Simulation, Volume I - Models, RAC-TP-266 (AD 387 510), Research Analysis Corporation, McLean, Virginia, Apr. 1967.

[60] Lazarsfeld, P., ed., Mathematical Thinking in the Social Sciences, The Free Press, New York, 1954.

[61] Louer, P. E., et al., Conceptual Design for the Army in the Field, Alternative Force Evaluation, CONAF Evaluation Model IV, OAD-CR-60, Part I - Model Description (AD 923 581L), Sept. 1974; Part II - Input Derivation (AD B000 570L), Nov. 1974; Part III - User's Handbook (AD B000 651L), Dec. 1974; General Research Corporation, McLean, Virginia.

[62] Low, L. J., "Concept and Plan" for the Theater-Level Gaming and Analysis Workshop, sponsored by the Office of Naval Research, Sep. 27-29, 1977.

[63] LULEJIAN--For model documentation, refer to: [64].

[64] Lulejian and Associates, Incorporated, The Lulejian - I Theater-Level Model (U), WSEG Report 259, Volume I - Overview of the Model (U), (AD C001 637L) CONFIDENTIAL; Supplement to Volume I - Test Data (U), (AD C001 638L) CONFIDENTIAL; Volume II - Model Logic and Equations (AD B003 814L); Volume III - Program Documentation (AD B003 815); Supplement to Volume III - Program Listing (AD B003 816L); Weapons Systems Evaluation Group, Arlington, Virginia, Oct. 1974.

[65] Miller, G., ed., The VECTOR-2 Theater-Level Combat Model, Volume I: User's Manual, VRI Report No. WSEG4,5 FR 77-1(R), Vector Research Incorporated, Ann Arbor, Michigan, Feb. 1977.

[66] Morse, P. M. and G. E. Kimball, Methods of Operations Research, OEG Report No. 54 (AD 200 782), Operations Evaluation Group, Office of the Chief of Naval Operations, Washington, D.C., 1946.

[67] Mulholland, R. P. and R. D. Specht, The Rate of Advance of the Front Line in Some World War II Campaigns, RM-1072, The RAND Corporation, Santa Monica, California, Apr. 1953.

[68] Newman, J. B., "Mathematics in Warfare by Frederick William Lanchester," The World of Mathematics: Volume 4, Simon and Schuster, New York, 1956, pp. 2138-2157.

[69] Office of Naval Research, Theater-Level Gaming and Analysis Workshop for Force Planning, Volume I - Proceedings, Sept. 22-29, 1977, L. J. Low, chm., Office of Naval Research, Arlington, Virginia, 1979.

[70] Ostermann, L. E., The Variation of Loss Ratios in Combat, Strategic and Tactical Studies, Lulejian and Associates, Incorporated, Redondo Beach, California, Dec. 1972.

[71] Overholt, J. L., Fitting Korean War Data by Statistical Methods, Professional Paper No. 32 (AD 705 349), The Center for Naval Analyses, Arlington, Virginia, May 1970.

[72] Overholt, J. L., et al., Study of Land/Air Trade-Offs, Volume IX: Statistical Analysis of Korean War Data--Spring 1951, Study No. 64 (AD 869 481), The Center for Naval Analyses, Arlington, Virginia, Apr. 1970.

[73] Quade, E. S., ed., Analysis for Military Decisions, R-387-PR, The RAND Corporation, Santa Monica, California, Nov. 1964.

[74] Quade, E. S. and W. I. Boucher, Systems Analysis and Policy Planning: Applications in Defense, R-439-PR (Abridged), The RAND Corporation, Santa Monica, California, Jun. 1968.

[75] Richardson, L. F., Statistics of Deadly Quarrels, Quadrangle Books, Chicago, Illinois, 1960.

[76] Samz, R. W., "Some Comments on Engel's 'A Verification of Lanchester's Law'," Operations Research 20, 1972, pp. 49-52.

[77] Shubik, M. and G. D. Brewer, Models, Simulations, and Games--A Survey, R-1060-ARPA/RC, The RAND Corporation, Santa Monica, California, May 1972.

[78] Steadman, R. C., Report to the Secretary of Defense on the National Military Command Structure, Requested by the President and conducted in the Department of Defense, Jul. 1978.

[79] Stockfisch, J. A., Models, Data, and War: A Critique of the Study of Conventional Forces, R-1526-PR, The RAND Corporation, Santa Monica, California, Mar. 1975.

[80] _____, Incentives and Information Quality in Defense Management, R-1827-ARPA (AD A036 914), The RAND Corporation, Santa Monica, California, Aug. 1976.

[81] Strauch, R. E., A Critical Assessment of Quantitative Methodology As a Policy Analysis Tool, P-5282, The RAND Corporation, Santa Monica, California, Aug. 1974.

[82] Sutherland, W. H., Some Basic Problems in War Gaming of Ground Combat, RAC-TP-153 (AD 470 102), Research Analysis Corporation, McLean, Virginia, Mar. 1965.

[83] Taylor, J. G., Attrition Modelling, Invited Presentation at the Third Systems Science Seminar, Delivered April 28, 1978 at Hochschule der Bundeswehr Munchen.

[84] _____, Lanchester-Type Models of Warfare, to be published by the Office of Naval Research, Arlington, Virginia, 1979.

[85] Taylor, V., "Subjectivity and Science: A Correspondence About Belief," Technology Review, Volume 81, Number 4, Feb. 1979, pp. 49-57.

[86] Thorp, K., et al., Comparison of the CARMONETTE Model with the TETAM Field Experiment, CAA-TP-75-8 (AD B008 161), U.S. Army Concepts Analysis Agency, Bethesda, Maryland, Dec. 1975.

[87] Tucker, S. A., ed., A Modern Design for Defense Decision: A McNamara-Hitch-Enthoven Anthology, Industrial College of the Armed Forces, Washington, D.C., 1966.

[88] U.S. Army Combat Developments Experimentation Command, Suppression Experiment (SUPEX), Final Report, CDEC Experiment FC029B, (AD B017 116L), Department of the Army, 1977.

[89] _____, Introductory pamphlet; U.S. Army CDEC Combat Developments Experimentation Command, Department of the Army, 1978.

[90] U.S. Army Concepts Analysis Agency, Firepower Potential Methodology Review (U), CAA-TP-73-7 (AD 529 179) CONFIDENTIAL, U.S. Army Concepts Analysis Agency, Bethesda, Maryland, Dec. 1973.

[91] U.S. General Accounting Office, Computer Simulations, War Gaming, and Contract Studies, B-163074, U.S. GAO, Washington, D.C. Feb. 1971.

[92] _____, Case Studies of Auditing in a Computer-Based Systems Environment, U.S. GAO, Washington, D.C., 1971.

[93] _____, Advantages and Limitations of Computer Simulation in Decisionmaking, B-163074, U.S. GAO, Washington, D.C., May 1973.

[94] _____, Improvement Needed in Documenting Computer Systems, B-115369, U.S. GAO, Washington, D.C., Oct. 1974.

[95] _____, Review of the 1974 Project Independence Evaluation System, OPA-76-20, U.S. GAO, Washington, D.C., Apr. 1976

[96] _____, Ways to Improve Management of Federally Funded Computerized Models, LCD-75-111, U.S. GAO, Washington, D.C., Aug. 1976.

[97] _____, An Evaluation of the Use of the Transfer Income Model--TRIM--To Analyze Welfare Programs, PAD-78-14, U.S. GAO, Washington, D.C., Nov. 1977.

[98] _____, Mission Budgeting: Discussion and Illustration of the Concept in Research and Development Programs, PSAD-77-124, U.S. GAO, Washington, D.C., Jul. 1977.

[99] _____, A Critique of the Performance of the Defense Systems Acquisition Review Council: Billions in Public Funds Involved, PSAD-78-14, U.S. GAO, Washington, D.C., Jan. 1978.

[100] _____, Models and Their Role in GAO, PAD-78-84, U.S. GAO, Washington, D.C., Oct. 1978.

[101] _____, Uses of National Economic Models by Federal Agencies, PAD-79-24, U.S. GAO, Washington, D.C., May 1979.

[102] VECTOR--For model documentation, refer to: [22] and [65].

[103] Vector Research Incorporated, "Target Acquisition Process," Unpublished slide presentation.

[104] Wainstein, L., Some Allied and German Casualty Rates in the European Theater of Operations, P-989 (AD 779 850), Institute for Defense Analyses, Arlington, Virginia, Dec. 1973.

[105] _____, Rates of Advance in Infantry Division Attacks in the Normandy-Northern France and Siegfried Line Campaigns, P-990 (AD 779 882) Institute for Defense Analyses, Arlington, Virginia, Dec. 1973.

[106] _____, An Examination of the Parsons and Hulse Papers on Rates of Advance, P-991 (AD 779 848), Institute for Defense Analyses, Arlington, Virginia, Dec. 1973.

[107] Walker, L. K., et al., Comparison and Evaluation of Four Theater-Level Models: CEM IV, IDAGAM I, LULEJIAN-1, VECTOR-1, WSEG Report 299, (AD A036 325), Weapons Systems Evaluation Group, Arlington, Virginia, Sept. 1976.

[108] Weiss, H. K., Requirements For a Theory of Combat, BRL Memorandum Report No. 667 (AD 13717), Ballistic Research Laboratories, Aberdeen Proving Ground, Maryland, Apr. 1953.

[109] _____, "Lanchester-Type Models of Warfare," Proceedings of the First International Conference on Operational Research, Operations Research Society of America, 1957, pp. 82-99.

[110] _____, "Combat Models and Historical Data: The U.S. Civil War," Operations Research 14, 1966, pp. 759-790.

[111] Whitson, W. L., et al., Preliminary Evaluation of Close Air Support Operations in Korea--and Appendixes, ORO-R-3(FEC) (ATI 144 065), Operations Research Office, The Johns Hopkins University, Chevy Chase, Maryland, Aug. 1951.

[112] Willard, D., Lanchester As Force in History: An Analysis of Land Battles of the Years 1618-1905, RAC-TP-74 (AD 297 375L), Research Analysis Corporation, McLean, Virginia, Nov. 1962.

[113] Zehna, P., Selected Methods and Models in Military Operations Research, Department of Operations Research, Naval Postgraduate School, Monterey, California, 1971, (Government Printing Office Stock No. 0851-0054, Washington, D.C.).

[114] Zimmerman, R. E., et al., Theater Battle Model (TBM-68): Volume VII - Technical Report, RAC-R-36 (AD 755 534), Research Analysis Corporation, McLean, Virginia, Jan. 1968.

[115] _____, RAC Presentation at DA "FEBA Movement" Seminar, (Unpublished text), Research Analysis Corporation, McLean, Virginia, Jan. 28, 1971.

APPENDIX IV

COMPTROLLER

ASSISTANT SECRETARY OF DEFENSE
WASHINGTON, D C 20301

24 SEP 1979

Mr. Harry S. Havens
Director, Program Analysis Division
United States General Accounting Office
Washington, D.C. 20548

Dear Mr. Havens:

This is in reply to your letter to the Secretary of Defense regarding your report dated August 3, 1979, on "Models, Data and War: A Critique of the Foundation for Defense Decisions," OSD Case #5248.

The GAO report examines the role of quantitative methodology in the analysis of public policy issues. The principal recommendations are that the Department of Defense should reassess its management practices with respect to policy assisting models employed in Defense Decisions and develop a plan that would allow the Joint Chiefs of Staff to serve as the Defense Establishment's principal analytic adviser on matters pertaining to the phenomenology of combat.

The Department of Defense feels that the present structure of the Department provides for satisfactory management of Defense studies. A Scientific and Technical Advisor is presently provided for within the structure of the Joint Chiefs of Staff. DoD Directive 5010.22 delineates the procedures to be followed in Defense studies to insure that models used are transparent, well documented and that all assumptions made are explicitly stated.

Enclosed, for your consideration, are more detailed comments prepared by the OSD in conjunction with the Joint Chiefs of Staff and the Services.

Sincerely,

Fred P. Wacker

Fred P. Wacker
Assistant Secretary of Defense

Enclosure

APPENDIX IV

GENERAL COMMENTS

The GAO report entitled "Models, Data and War: A Critique of the Foundation for Defense Decision" examines the role of quantitative methodology in the analysis of public policy issues. Chapter two is good in its treatment of the current theories of modeling and the role models play when incorporated into policy studies. The report is correct when it emphasizes that models, to the extent possible, should be transparent and that all assumptions employed in the model should be open and explicit. As the author accurately points out, models should be looked upon not as "'providers of solutions', but rather as framework which permit science and the judgment of experts in numerous subfields to be brought together - made explicit - and utilized to enhance and extend a decision maker's judgment." However, the report does have some serious drawbacks. Although many recommendations are made to improve the "current design and management of quantitative tools," the report fails to support these.

NOTE A*

SPECIFIC COMMENTS

- Scope of the Report

The report is exceptionally narrow in scope in that it deals exclusively with dynamic theater level models to the exclusion of the more common static models. Although the static model does have some limitations it suffers from few of the same maladies pointed out in the report with respect to dynamic models. Generally, the static model is transparent and well documented with explicit statements concerning assumptions.

NOTE B*

The author's apparent belief that decision makers at the highest level of the Defense Department depend heavily on the results of specific models when making decisions is erroneous. Models are employed within the framework of studies, with the intimate workings of the model seldom going beyond the study director. The results of the model are interpreted and evaluated with relationship to the study. The degree to which the results of the model are incorporated in the study report is usually dependent upon the study director's judgment of the validity of the model and the confidence that he has in the results. Models in themselves seldom enter into the higher echelons of decision making; studies do.

pp. 80-82*

- Role of the Joint Chiefs of Staff

The report recommends that plans and recommendations be prepared ". . .which would enable the Organization of the Joint Chiefs of Staff to serve as the Defense Establishment's principal analytic adviser on matters pertaining to the phenomenology of combat." In addition, it recommends that it provide the Defense analytic community with a senior adviser.

pp. 105-109*

*Page numbers refer to report citation; GAO notes follow the Department's letter.

pp. 105-109*

The Studies, Analysis and Gaming Agency (SAGA) within the JCS presently performs analyses of alternative military strategies, force structures and postures in support of the responsibilities of the JCS to the Secretary of Defense. In this context they are the principal analytic advisers on matters pertaining to the "phenomenology of combat".

The JCS already has a senior scientific and technical adviser within the SAGA organization. It is his responsibility to advise members of the JCS on issues of a scientific and technical nature. Since each DoD component has responsibilities for defense analyses that are not within the purview of the JCS, it would not be practical to establish an institutional focus for all Defense analysis within that organization.

pp. 80-86*

- Reference Data Base

Measures toward ". . .establishing a reference data base to support multiple users" are already being implemented. The Office of the Assistant Secretary of Defense (Program Analysis and Evaluation) is presently updating the Defense Force Planning Data Base. This will assist in solving many of the problems associated with data collection efforts identified in the report. It will provide users with a consistent source for input data.

- Prescribing Models and Data to be used

Presently the Department of Defense does specify what models and what data are to be used in specific studies. However, this technique is not, and should not, be used in every study. In many cases there is no standard, prescribed method of solution. In such cases the DoD component conducting the study should have the responsibility and freedom to develop the necessary methodology, models, and data to satisfy its requirement.

- Management and Conduct of Studies and Analysis

The report dwells heavily on the management aspects of studies and analysis. Specifically, it recommends "an open explicit understanding of the assumptions underlying a study's conclusions" be required, and identification of ". . .corrective measures to insure that models are well documented, comprehensively appraised and updated on a formal basis."

Department of Defense Directive 5010.22 specifies the policy to be followed with respect to studies and analysis. Specifically, it requires that a report associated with each study document all information on assumptions made, models used, model modifications, effectiveness measures, criteria and sources of intelligence or data used.

*Page numbers refer to report citation; GAO notes follow the Department's letter.

pp. 105-106*

OSD is in the process of establishing a formalized procedure for spot auditing of study efforts throughout DoD, concentrating on several mission areas each year. This overall review of the content, methodologies and use of studies and analysis will give close attention to modeling and other analytical tools employed.

NOTE C*

- Theater Level Models

The report's enthusiastic acceptance of the Vector II theater model may be premature. The assumption made is that because Vector II is more detailed it is necessarily better. Vector II at this point is an unproven quantity. Separate specialized models can best embody the engineering detail of systems that too often is not adequately represented in other theater level models. Both the Army and the Air Force employ the hierarchical technique insuring adequate detail within the model at the lowest levels. Up to theater level models can be constructed using a set of separate but consistent and related combat models. A shift toward an exclusive use of theater level models would not be constructive.

- Quantitative Methodology vs Professional Judgment

NOTE D*

It is quite clear that the author recognizes that "squishy problems" are not modeled purely using hard scientific data. What is unclear is whether he contends that because policy assisting models ". . .do not produce 'objective', or verifiable solutions" they are inferior to the intuitive judgment of the decision maker. Regardless of how "squishy" the policy problem, the properly managed, professionally prepared study can assist in the decision making process.

- Congressional Recommendations

NOTE E*

The report recommends that Congress should "require an open, explicit understanding of the assumptions underlying a study's conclusions or knowledge of the identity of the decision makers involved in the study, their background experience and institutional affiliations. . .." The fundamental principle of any analysis is that it should stand solely on its own merits, independent of the reputation and institutional affiliation of the decision makers. The report implies that the individual, not the analysis, should be the primary concern of Congress.

In conclusion we feel that the report generally does an excellent job of identifying those deficiencies in modeling previously identified by the Department of Defense. We feel that the existing structure within DoD can and does adequately manage Defense studies. Recommendations toward establishing another organization within the existing structure are not constructive. The process of model improvement is an evolutionary process, not a revolutionary one. Working toward developing "better" models continues to be a prime objective of the modeling community. Current Defense policy supports this objective.

***Page numbers refer to report citation; GAO notes follow the Department's letter.**

GAO NOTES

A. "Although many recommendations are made to improve the 'current design and management of quantitative tools,' the report fails to support these."

GAO's report intends no appraisal of specific models and studies, nor any criticism of specific decisions. It does, however, cite a few examples which we believe are representative of the Department's management of models (pp. 73-75) and data (p. 75). Let us say very simply that this list includes a specific example which has contributed to suboptimizations in, and methodological discontinuities between, analyses supporting requirements determination, force planning, and budgeting. We would expect any one of these conditions to support a recommendation for improvement. 1/

B. "The report is exceptionally narrow in scope in that it deals exclusively with dynamic theater level models to the exclusion of the more common static models."

As pointed out in the footnote on page 52, the Services use many different types of combat models for analyses. Such models are "dynamic" in the sense that they attempt to represent the space-time and environmental consequences of real world combat. As a general rule, each model is specific to one level of combat interaction. Collectively, the models span a spectrum of combat interactions--ranging from one-on-one duels at one end, to theater-level combat at the other. Their use includes analyses relating to weapon system, force mix, force level, and contingency planning decisions.

DOD is correct in pointing out that the report deals exclusively with theater-level combat models. This choice provides a simple illustration of how four models render profoundly different interpretations of the same real world phenomena (see ch. 4). Explicitly, we underscore the need for model transparency--a feature which is by no means unique to the theater-level models. Implicitly, we focus on the importance of understanding

---

1/For further discussion on the need for improvement, see: Hardison, D. C., study director, Review of Army Analysis, Volume I - Main Report, Special Study Group, Department of the Army, Washington, D.C., Apr. 1979; Office of Naval Research [69]; and Steadman [78].

the phenomenology of combat. It should be apparent that if attrition and the synergistic effects of lower-level combat were truly understood, there would not be such a disparity in the modeling of theater-level attrition. Conversely, to the extent that these things are not well understood--the analytical basis for assessing weapon system effectiveness appears to warrant continuing attention.

Static models--more appropriately termed static indices or indicators--are widely used within OSD and by the JCS for balance assessments. However, the static indicators are not combat models; they are "bean counts" devoted to comparisons of things--people, weapons, units, etc. The more sophisticated variations try to account for all the firepower capability in different units, in effect comparing unlike systems to each other. Among these are "judgmental" firepower scores, associated with the terms Weapon Effectiveness Indices and Weapon Unit Value (WEI/WUV); and "laboratory" scores, associated with the term firepower potential (FPP). The "transparency" of a static indicator usually depends upon its "sophistication." 1/ Because they shed little light on the phenomenology of combat, we have omitted their discussion in this report. As pointed out in the MEFORD study:

> "The principal objection to all of the measures generically described as static indicators is that they do not provide an adequate appreciation of the relevant dimensions of the force planning problem and thus do not provide adequate appraisals of the consequence of the choice of one alternative force level over another. They do present to decision makers a useful anatomical description of the extent to which his forces have improved or deteriorated relative to those of the putative enemy. But these anatomical descriptions provide an insufficient basis for sound year-to-year

---

1/For a brief discussion of firepower potential, see pp. 54-56. For a critical review of the limitations of static indices: see Bode [7]; Dondero [25]; and the Congressional Budget Office, Assessing the NATO/Warsaw Pact Military Balance, U.S. Government Printing Office, Washington, D.C., Dec. 1977.

force planning. Important issues which cannot be settled from a review of the inventory-type descriptors include the following:

"1. The effect of alternative assumptions about the timing of mobilization and deployment in relation to the actual outbreak of hostilities.

"2. The effects on the progress of battle of the possible alternative deployments or tactics of available troops, for example, linear defenses versus highly mobile reserves, or concentration and breakthrough versus advances along a wide front.

"3. The relationship of the available troop units to operating space, both linearly along the front, and in depth throughout the area of operations, as it may expand or contract.

"4. The effects of ground lost or gained as this may affect the deployment integrity of forces, the practical scheme of maneuver, and the possible reorientation and the capabilities of LOCs [lines of communications].

"5. The dynamic effects of interdiction campaigns on friendly and enemy LOCs.

"6. The assessment of assumptions of reasonable variability in the estimated quality and endurance of opposed combat forces." 1/

C. "The report's enthusiastic acceptance of the Vector II theater model may be premature. . . . Both the Army and the Air Force employ the hierarchical technique . . . Up to theater level models can be constructed using a set of separate but consistent and related combat models."

1/See Dondero [25], pp. 49-50.

The Department misrepresents the facts when it asserts: (1) that the report enthusiastically accepts VECTOR-2; and (2) that "the assumption made is that because Vector II is more detailed it is necessarily better." Had this been the intent, the report would have so stated. On the other hand, it should come as little surprise that VECTOR fares well in a technical comparison. The Department's "principal analytic advisers on matters pertaining to the phenomenology of combat" are devoting substantial resources to bring it "on-line". We presume their faith is not misplaced; and that VECTOR-2 will become one of a variety of useful models within the Department.

The Department's second point extols the virtues of a "hierarchy of models" (for a description of the technique, see fn. 1, p. 79). We readily support this concept. But a concept is one thing; and its in-practice realization, quite another.

-- The Air Force presently has an operational hierarchy extending up to theater-level air warfare. This hierarchy includes detailed sub-models focusing on such diverse aspects of the air war as: (1) the scheduling of aircraft maintenance; (2) airfield defense; and (3) many-on-many aircraft engagements in an electronic environment. We have been told that the flexibility of this hierarchical arrangement has proven extremely useful in supporting Air Force studies and analyses. The Air Force's models are omitted from this report's discussion because they provide only a rudimentary representation of the ground battle.

-- The Army has employed the hierarchical concept on a limited scale since the early 1970s; but only to the extent of linking a computer simulation of battalion combat to a division-level war game. The feasibility of adopting a hierarchy of models extending from individual weapons to theater-level combat is now under study (fn. 1, p. 79).

In this perspective, the Department has pointed out that: "A shift toward an exclusive use of theater level models would not be constructive." We readily agree. As we noted on p. 72, "Models are tools"; and there is an appealing simplicity in using a tool suited to the task at hand. The operational imperative is that the model

be part of a consistent analytical framework (again, fn. 1, p. 79).

D. "What is unclear is whether he contends that because policy assisting models. . . are inferior to the intuitive judgment of the decisionmaker."

As we stated on page 6: "Policy assisting models are one of the most significant decisionmaking tools of our day* * * . They should be looked upon not as a 'provider of solutions,' but rather as a framework which permits science and the judgment of experts in numerous subfields to be brought together--made explicit --and utilized to enhance and extend a decisionmaker's judgment." We wholeheartedly agree with the Department's statement that: "Regardless of how 'squishy' the policy problem, the properly managed, professionally prepared study can assist in the decision making process." Insuring such studies--improving existing practices--is the purpose of our recommendations.

E. "The fundamental principle of any analysis is that it should stand solely on its own merits, independent of the reputation and institutional affiliation of the decision makers. The report implies that the individual, not the analysis, should be the primary concern of Congress."

We believe it is important neither to denigrate, nor to overstate, the credibility of analyses offered for congressional decisionmaking. For that reason, the report recommends "an open, explicit understanding * * * or knowledge of * * *" (see p. 113). It would be a pitfall to believe that the results of complex analytical studies, embodying a substantial volume of intuitive judgments, can stand solely on their own merits. As the Department pointed out earlier in its comments:

> "Models are employed within the framework of studies, with the intimate workings of the model seldom going beyond the study director. The results of the model are interpreted and evaluated with relationship to the study. The degree to which the results of the model are incorporated in the study report is usually dependent upon the study director's judgment of the validity of the model and the confidence that he has in the results."

The magnitude of these judgments can well exceed the practicalities of delineation. Indeed,

> "It is impossible to combine given data and given measures of effectiveness into a model without setting up new patterns of interaction, whether or not this is explicitly reported. * * * A case in point is the use of both firepower scores and force ratio advance data. These interact * * * [to] create a new measure which can be described loosely in ways such as 'weapons numbers times lethality equals force ratios relative to movement over time' but there is little data available to anyone how many meters twelve 155mm howitzers advance the FEBA in hours which is one explicit result of this new 'measure of effectiveness'."1/

This is not to say that implicit judgments--as in this linkage between howitzers and frontline movement; or perhaps in an analyst's trade-off between casualties and ground gained--invalidate the analysis. It is to say that the analyst's impact on an analysis is not an arbitrary one.

The Department has said our "report implies that the individual, not the analysis, should be the primary concern of Congress." We would prefer to think the message is more towards safeguarding the integrity of analysis in support of public policy.

The essence of the scientific method is that it is replicable. When properly documented, the analysis of a rigorously quantifiable problem can be verified or refuted by an independent critic--solely on the grounds of the mathematical analysis within the study and the empirical connection between the analysis and the substantive problem. This is more difficult in the analysis of a squishy problem. In general, no objective standards exist by which the linkages between analysis and the substantive problem can be validated.

At issue, therefore, are the quality assurance and credibility of analyses offered for decisionmaking. When

---

1/See Cordesman [23], p. 186.

it becomes necessary to support Departmental positions with quantitative studies, those studies should provide a suitable framework for congressional review. If--as in the analysis of many squishy problems--the study cannot stand alone, we believe the Congress should have knowledge of, and access to, the responsible analyst.

(973520)

* U.S. GOVERNMENT PRINTING OFFICE : 1980-620-306/00